القياس النفسي

في ظل النظرية التقليدية والنظرية الحديثة

الدكتور
ابراهيم محمد محاسنة

دار جرير
للنشر والتوزيع



mohamed khatab



دار جرير
للنشر والتوزيع
www.darjareer.com





القيـاس النفسي

في ظل النظرية التقليدية والنظرية الحديثة

القيــاس النفســي

في ظل النظرية التقليدية والنظرية الحديثة

الدكتور

إبراهيم محمد محاسنة

الطبعة الأولى

١٤٣٤ هـ - ٢٠١٣م



القياس النفسي في ظل النظرية التقليدية والنظرية الحديثة

د. ابراهيم محمد المحاسنة

رقم الإيداع لدى دائرة المكتبة الوطنية (٢٠١٢/٩/٣٣٢٩)

رقم التصنيف : ١٥٥.٢٨

الواصفات:/علم النفس//القياس (منطق)//

الطبعة الأولى ١٤٣٤هـ - ٢٠١٣م

حقوق الطبع محفوظة للناشر
All rights reserved



عمّان - شارع الملك حسين - مقابل مجمع الفحيص التجاري
هاتف : ٤٦٥١٦٥٠ - فاكس : ٤٦٤٣١٠٥ - ٦ - ٠٠٩٦٢
ص . ب . : ٣٦٧ عمّان ١١١١٨ الأردن
E-mail: dar_jareer@hotmail.com

ردمك ISBN 978- 9957 – 38 – 262-9

جميع حقوق الملكية الفكرية محفوظة لدار جرير للنشر والتوزيع عمان- الأردن ويحظر طبع أو تصوير أو ترجمة أو إعادة تنضيد الكتاب كاملا أو مجزأ أو تسجيله على أشرطة كاسيت أو إدخاله على الكمبيوتر أو وضعه على مواقع الكترونية أو برمجته على اسطوانات ضوئية إلا بموافقة الناشر خطيا.

الإهداء

إلى والدي العزيزين

إلى زوجتي ورفيقة دربي

إلى فلذات كبدي

تامر

تمارى

سارة

رؤى

والذي لم يأت بعد

د. إبراهيم محمد عبدالله المحاسنة

الفهرس

الإهداء ٥

الباب الأول: النشأة والمفاهيم الأساسية

الفصل الأول: نشأة القياس ١٩

نشأة القياس ٢١

نشأة وتطور نظرية القياس ٢٣

القياس لدى المسلمين ٢٦

القياس في العصور الحديثة ٢٨

الفصل الثاني: مفاهيم القياس ٤٩

علم القياس ٥٠

مفاهيم أساسية في القياس ٥١

الاختبار ٥١

التقويم ٥١

الاختيار ٥٢

التصنيف والتعين ٥٢

الإرشاد ٥٢

القياس ٥٢

مستويات القياس ٥٣

المقياس الاسمي ٥٤

المقياس الرتبي ٥٥

المقياس الفتري ٥٥

المقياس النسبي ٥٦

المتغير والثابت ٥٨

مفاهيم أساسية في الإحصاء ٥٩

الإحصاء الوصفي ٥٩

مقاييس النزعة المركزية ٦٠
الوسط الحسابي ٦٠
المنوال ٦٠
الوسيط ٦٠
ثانيا: مقاييس التشتت ٦١
المدى ٦١
التباين ٦١
الانحراف المعياري ٦٢
الانحراف المتوسط ٦٢
الإحصاء الاستنتاجي ٦٣
معامل الارتباط ٦٤
معامل ارتباط بيرسون ٦٦
معامل الارتباط الرتبي ٧١
تفسير معامل الارتباط ٧٣
التباين المفسر ٧٣
الخطأ المعياري للتقدير ٧٤
معادلة التنبؤ ٧٤
العوامل المؤثرة في معامل الارتباط ٧٧
ضيق المدى ٧٨
دمج المجموعات ٧٩
صور أخرى لمعامل الارتباط ٨٠
معامل الارتباط فاي (Φ) ٨٠
تفسير معامل الارتباط فاي Φ ٨١
معامل الارتباط الرباعي ٨٤
معامل الارتباط ثنائي التسلسل ٨٧
معامل الارتباط بايسيريال ٨٩

الباب الثاني : النظرية التقليدية في القياس

الفصل الأول: النظرية التقليدية في القياس..................... ٩٥

افتراضات النظرية التقليدية ١٠١

استنتاجات خاصة بافتراضات النظرية التقليدية ١٠٨

الفصل الثاني: مفهوم الثبات..................... ١١٩

طرق حساب الثبات ١٢٦

طريقة تطبيق وإعادة تطبيق الاختبار..................... ١٢٧

طريقة الصور المتكافئة..................... ١٢٨

مفهوم التوازي..................... ١٢٩

طريقة التجزئة النصفية ١٣٠

طريقة الاتساق الداخلي..................... ١٣٤

الحالة العامة طريقة سبيرمان – براون ١٣٥

معادلة كرونباخ الفا ١٣٦

معادلة كودر – ريتشاردسون ٢٠ .... ١٣٦

معادلة كودر – ريتشاردون ٢١ .... ١٣٧

ثبات الاختبار محكي المرجع ١٣٩

معادلة ليفنجستون ١٣٩

معادلة كابا ،سوامنيثان،هاميلتون،الجانيا ١٤٠

معادلة كابا المعدلة ١٤١

معادلة سابكوفياك ١٤٢

العوامل المؤثرة على الثبات ١٤٣

كيف يؤثر طول الاختبار على الثبات ١٤٤

أمثلة توضيحية لمفهوم الثبات..................... ١٤٦

الفصل الثالث : مفهوم الصدق ١٥١

أنواع الصدق..................... ١٥٢

صدق المحتوى ١٥٢
الصدق الظاهري ١٥٢
الصدق المنطقي (العيني) ١٥٣
الصدق المرتبط بالمحك ١٥٤
الصدق التنبؤي ١٥٥
الصدق التلازمي ١٥٦
صدق المفهوم (البناء) ١٥٧
صدق الاختبار محكي المرجع ١٥٩
الصدق الوصفي ١٦٠
الصدق الوظيفي ١٦٠
صدق الانتقاء للنطاق السلوكي للاختبار ١٦٠
مؤشرات الصدق ١٦١
مؤشرات صدق المحتوى ١٦١
مؤشرات الصدق الظاهري 162
مؤشرات الصدق المرتبط بمحك ١٦٢
مؤشرات الصدق التلازمي ١٦٣
مؤشرات الصدق التنبؤي ١٦٣
مؤشرات صدق البناء ١٦٣
العوامل المؤثرة في الصدق ١٦٤
عوامل خاصة بالاختبار ١٦٤
عوامل خاصة بتطبيق وتصحيح الاختبار ١٦٤
عوامل خاصة بالمفحوص ١٦٤
العلاقة بين صدق الاختبار وثباته ١٦٥

الباب الثالث نظرية السمات الكامنة

الفصل الأول: الأطر النظرية.......................................١٦٩
منحنى خصائص الفقرة.......................................١٧١
معلمة صعوبة الفقرة.......................................١٧٢
معلمة تمييز الفقرة.......................................١٧٢
خلاصة.......................................١٨١
الفصل الثاني: نماذج منحنى خصائص الفقرة.......................................١٨٣
الاقتران اللوغريتمي.......................................١٨٤
النموذج اللوغريتمي أحادي المعلمة.......................................١٨٤
النموذج اللوغريتمي ثنائي المعلمة.......................................١٨٨
النموذج ثلاثي المعلمات.......................................١٩١
ظاهرة التمييز السالب.......................................١٩٧
خلاصة.......................................٢٠١
الفصل الثالث: تقدير معلمات الفقرة.......................................٢٠٣
المجموعة (الثابتة) المشتركة لمعلمات الفقرة.......................................٢٠٨
خلاصة.......................................٢١٣
الفصل الرابع: منحنى خصائص الاختبار.......................................٢١٥
خلاصة.......................................٢٢٥
الفصل الخامس: تقدير قدرة المفحوصين.......................................٢٢٧
إجراءات تقدير القدرة.......................................٢٢٨
علاقة ثبات الفقرة بتقدير قدرة المفحوصين.......................................٢٤١
خلاصة.......................................٢٤٤

توزيع القدرة المقدرة ٢٤٤
علاقة ثبات الفقرة بتقدير قدرة المفحوصين.... ٢٤٤
ثبات قدرة المفحوصين ٢٤٥
الفصل السادس:اقتران المعلومات.... ٢٤٧
اقتران معلومات الفقرة ٢٥٠
اقتران معلومات الاختبار ٢٥١
تعريف اقتران معلومات الفقرة.... ٢٥٣
حساب اقتران معلومات الاختبار.... ٢٦٣
تفسير اقتران معلومات الاختبار.... ٢٦٩
خلاصة ٢٧١
الفصل السابع : تدريج الاختبار وتحديد خصائصه ٢٧٣
إجراءات تدريج الاختبار ٢٧٤
مشكلة وحدة القياس.... ٢٧٥
تدريج الاختبار في ظل النموذج أحادي المعلمات (نموذج راش) ٢٧٦
ملخص إجراءات تدريج الاختبار.... ٢٨٣
خلاصة ٢٨٥
تحديد خصائص الاختبار ٢٨٧
تطوير اختبار بالتدريج المسبق ٢٨٨
الأهداف العملية للاختبار.... ٢٨٩
أنواع الاختبارات حسب النظرية الحديثة.... ٢٩٠
اختبارات التصفية.... ٢٩٠
الاختبارات واسعة النطاق.... ٢٩٠

اختبارات الفهم (الفجوات) ... ٢٩١

الباب الرابع: مقارنة بين النظرية الكلاسيكية والنظرية الحديثة للقياس

الفصل الأول: النظريات والنماذج ... ٢٩٥

أهمية نظريات القياس ونماذجها ... ٢٩٦

مفهوم الثبات مقابل مفهوم اقتران المعلومات ... ٢٩٩

مجموعة التطبيق مقابل مقياس القدرة ... ٢٩٩

مقبولية مؤشرات الفقرات مقابل تطابق الفقرات مع النموذج ... ٣٠٠

الفصل الثاني النظرية الكلاسيكية ... ٣٠١

محددات النظرة الكلاسيكية ... ٣٠٢

أوجه القصور ... ٣٠٣

ميزات النظرية الكلاسيكية ... ٣٠٦

تحليل الفقرات حسب النظرية الكلاسيكية ... ٣٠٦

اختيار الفقرات ... ٣٠٧

تطبيقات النظرية السمات ... ٣٠٨

تحليل الفقرات حسب النظرية الحديثة ... ٣٠٩

اختيار الفقرات ... ٣٠٩

قائمة المراجع العربية ... ٣١٥

قائمة المراجع الأجنبية ... ٣١٩

فهرس الجداول

جدول (١) مستويات القياس وخصائصها والتي توضح هرمية المقاييس ٥٤
جدول (٢) نوع وقوة الارتباط بين المتغيرات. ٦٥
جدول:(٣) مكونات الدرجة الحقيقية حسب النظرية التقليدية ١٠٦
جدول:(٤) مستويات معلمتي الصعوبة والتمييز حسب النظرية الحديثة. ١٧٧
جدول:(٥) احتمال الإجابة حسب الصعوبة والتميز حسب النموذج الأحادي ١٨٧
جدول:(٦) احتمال الإجابة الصحيحة لفقرة صعوبتها (١) وتمييزها(٠.٥). ... ١٩٠
جدول.(٧) قيم احتمال الإجابة الصحيحة على فقرة حسب النموذج ثلاثي. ..... .١٩٦
جدول:(٨) مستويات وأوصاف معلمة التميز. ١٩٩
جدول:(٩) القيم الخاصة بتقدير قدرة المفحوصين في الفترة الأولى. ٢٣٢
جدول:(١٠) القيم الخاصة بتقدير قدرة المفحوصين في الفترة الثانية ... .٢٣٦
جدول:(١١) القيم الخاصة بتقدير قدرة المفحوصين في الفترة الثالثة... . . .. ٢٣٩
جدول:(١٢) قيم المعلمات والقيم اللازمة لحساب اقتران معلومات الفقرة ٢٥٥
جدول:(١٣) قيم المعلمات والقيم اللازمة لحساب اقتران معلومات الفقرة ... .٢٥٨
جدول:(١٤) قيم المعلمات والقيم اللازمة لحساب اقتران معلومات الفقرة ٢٦٢
جدول:(١٥) قيم المعلمات والقيم اللازمة لحساب اقتران معلومات الاختبار . .٢٦٨
جدول:(١٦) استجابات (١٦) مفحوص على اختبار مكون من (١٠) فقرات. . ..٢٦٨
جدول:(١٧)علامات المفحوصين والفقرات حسب إجراءات نموذج راش. ٢٧٨
جدول:(١٨) تقديرات معلمة الصعوبة للفقرات الواردة في جدول (١٦). ٢٨٠
جدول:(١٩) العلامات الخام وتقديراتها من خلال بيانات الجدول (١٦). ٢٨٢
جدول:(٢٠) مقارنة بين النظرية التقليدية ونظرية السمات الكامنة ٣١١

فهرس الأشكال

شكل (١): التحويلات الخطية التي يسمح بها في مستوى القياس الفتوي ٥٨
شكل (٢): التوزيع الطبيعي ٦٤
شكل (٣) :قوة العلاقة الخطية بين متغيرين ٦٩
شكل (٤) قوة واتجاه العلاقة (الارتباط) من خلال شكل الانتشار............٧٥
شكل (٥- أ) درجات تقع ضمن مدى ضيق..٧٨
شكل (٥- ب) درجات تقع ضمن أكثر اتساعا..٧٨
شكل رقم (٦أ – ٦ ب) ظاهرة دمج المجموعات وأثرها على الارتباط؟ ٧٩
شكل (٧) الارتباط بين الدرجات الملاحظة والحقيقية وأثرها على الثبات.........١٢١
شكل (٨)الارتباط الدرجات الملاحظة والحقيقية ١٢٢
شكل (٩) توضيح لمعامل الارتباط بين الدرجات.....................................١٢٣
شكل (١٠) منحنى خصائص لفقرة مقبولة منطقيا.......................................١٧٢
شكل(١١أ) منحنى الخصائص لفقرة سهلة وتمييزها متدني ١٧٣
شكل(١١ب) منحنى الخصائص لفقرة متوسطة الصعوبة وتمييزها مقبول...........١٧٣
شكل(١١ج) منحنى الخصائص لفقرة متوسطة الصعوبة وتمييزها مقبول.............١٧٤
شكل(١٢ أ) منحنى خصائص فقرة تمييزها عالي وصعوبة أعلى من المتوسطة.......١٧٥
شكل(١٢ب) منحنى خصائص فقرة ذات تميز متوسط وصعوبة متوسطة.......... ١٧٥
شكل(١٢ج) منحنى خصائص فقرة ذات تميز متوسط وصعوبة متوسطة...........١٧٦
شكل(١٢ د) منحنى الخصائص لفقرة ذات تميز تام وصعوبة عالية...............١٧٦
شكل (١٣أ) منحنى الخصائص لفقرة متوسطة الصعوبة والقدرة التمييزية ١٧٧
شكل (١٣ ب) منحنى الخصائص لفقرة سهلة ومتدنية القدرة التمييزية ١٧٨
شكل (١٣ ج) منحنى خصائص أعلى من متوسطة الصعوبة وتمييزها متوسطة.....١٧٨

شكل (١٣ د) منحنى الخصائص لفقرة صعبة ومتدنية القدرة التمييزية ١٧٩
شكل (١٣ هـ) منحنى الخصائص لفقرة صعبة وتامة القدرة التمييزية ١٨٠
شكل (١٤) منحنى الخصائص للفقرة الواردة في المثال رقم (١) ١٨٧
شكل (١٥) منحنى خصائص الفقرة الواردة في المثال (٢) ... .١٩١
شكل (١٦) شكل منحنى الخصائص للفقرة الواردة في المثال (٣). .. .. .١٩٦
شكل (١٧) منحنى خصائص لفقرة ذات تميز سالب . . .. . . ١٩٧.
شكل (١٨أ) منحنى خصائص فقرة تميزها سالب في حال تكون الإجابة خاطئة . ..١٩٨
شكل (١٨ب) منحنى الخصائص لفقرة تميزها موجب في حال تكون
الإجابة صحيحة ١٩٨
شكل (١٩) شكل الاحتمال الملاحظ للإجابة الصحيحة على فقرة ٢٠٤
شكل (٢٠) منحنى الخصائص لفقرة مطابقة الاحتمالات الملاحظة
للإجابة الصحيحة ٢٠٦
شكل(٢١) شكل احتمالات الإجابة الصحيحة لمجموعتين مختلفتين ٢٠٨
شكل (٢٢) منحنى خصائص الفقرة المتطابق مع بيانات المجموعة الأولى.......... ..٢٠٩
شكل (٢٣) الاحتمالات الملاحظة للإجابة الصحيحة للمجموعة الثانية .... .٢٠٩
شكل (٢٤) منحنى خصائص الفقرة المتطابق مع بيانات المجموعة الثانية............ .٢١٠
شكل (٢٥) منحنى خصائص الفقرة المتطابق مع بيانات المجموعتين معا ٢١١
شكل (٢٦) منحنى خصائص الفقرة الأولى ٢١٧
شكل (٢٧) منحنى خصائص الفقرة الثانية ٢١٨
شكل (٢٨) منحنى خصائص الفقرة الثالثة ٢١٩
شكل (٢٩) منحنى خصائص الفقرة الرابعة ٢٢٠
شكل (٣٠) منحنى الخصائص لاختبار يتكون من ١٠ فقرات ٢٢١

شكل (٣١) منحنى اقتران المعلومات للفقرة.... .. . .٢٤٨
شكل (٣٢) منحنى اقتران المعلومات لفقرة ما.. .. . .٢٥٠
شكل (٣٣) منحنى اقتران الاختبار......... .. .. .٢٥٢
شكل(٣٤) شكل منحنى اقتران المعلومات لاختبار مكون من خمس فقرات ٢٦٩

الباب الأول

النشأة والمفاهيم الأساسية

الفصل الأول

نشأة القياس

مقدمة:

أجمع المعلمون والمهتمون في قضايا التربية والتعليم على أن القياس والتقويم حجر الأساس في عملية التطوير والتحديث والتجديد لما يشكله من أهمية بالنسبة للمعلم والطالب معاً ، وتشكل المرحلة الأخيرة من عملية التعلم والتعليم نقطة البداية لتعلم جديد أو لا حق ، وتهدف هذه العملية إلى معرفة مواطن الضعف والقوة في عمليتي التعلم والتعليم بهدف إدخال تحسينات عليها ، من حيث أساليب التدريس ، أو الوضع التعليمي ، أو المادة الدراسية وغير ذلك

كما يعتبر القياس ركناً أساسياً، وعنصراً هاماً من عناصر العملية التربوية بشكل عام، والعملية التدريسية بشكل خاص، ولا يستطيع المعلم في مدرسته والمدرس في جامعته أو كليته القيام بدوره الأساسي كمقوم بدون توفر الحد الأدنى من المعلومات والمهارات الأساسية في مجال القياس والتقويم بشكل عام، والاختبارات التحصيلية بشكل خاص، ولذلك يبدو الاهتمام واضحاً من قبل متخذي القرارات بتأهيل المعلمين في هذا المجال قبل الخدمة وأثنائها، وبتأهيل المدرسين في الجامعات، من خلال برامج موجهة لهذا الغرض.

وان المستعرض للأدب الخاص بالقياس والتقويم وخاصة في المجال التربوي يجد انه لم يخل أي نظام تربوي في أي عصر من العصور من ممارسات وإجراءات القياس باعتبارها ضرورة إنسانية للحياة البشرية ، فقد عرف الإنسان منذ القدم تقدير مواعيد الفصول مثل بداية الشتاء والصيف وإزهار النبات ومواعيد الزراعة و قطف الثمار وسموا المجموعات حسب إعدادها كالسرب والرهط والفوج وكذلك عدد السنوات والأشهر والأيام ولذلك يعتبر التأريخ لمواعيد الميلاد والوفاة شكل من أشكال القياس، كما سميت الأشياء والأفراد بخصائصها فقيل الشمس الحارة أو السنة المطيرة أي كثيرة المطر أو الشخص السريع البديهة أو الذكي أو المتحمس وهذه كلها تعتبر

ممارسات لعلم القياس مع الأخذ بعين الاعتبار اختلاف مستوياتها ، كما عرف الإنسان الاختبارات واستخداماتها حيث اعتبرها الصينيون معيارا للالتحاق بالوظائف ونقل عنهم الأوربيون هذا النظام وعرفوا الاختبارات الشفوية التي ظلت سائدة حتى أواسط القرن التاسع عشر في أوروبا، ثم دخلت عمليات وإجراءات القياس مرحلة جديدة في مطلع القرن العشرين بدخول اختبارات الذكاء على يد الفرد (Benet & Simon) بنية وسايمون، ويعتبر القياس احد اهم الأركان الأساسية للحياة البشرية، فهو حجر الزاوية لإجراء أي عملية أو قرار بهدف جمع المعلومات لأغراض التغيير والتعديل والتحسين في عمليات وممارسات الإنسان في الحياة اليومية تحسين عملية التعليم والتعلم في أي بلد ، كما ينظر للتقويم – وهو المفهوم الملازم للقياس – من قبل متخذي القرارات التربوية على مختلف المستويات على أنه الدافع الرئيس الذي يقود العاملين في أي مؤسسة على اختلاف مواقعهم في السلم الإداري إلى العمل على تحسين أدائهم وممارساتهم وبالتالي مخرجاتهم فالقياس يسهم في معرفة درجة تحقق الأهداف الخاصة بالعملية الاقتصادية أو التعليمية أو الصحية ، ويسهم في الحكم على صحة ودقة الإجراءات والممارسات المتبعة ، ويوفر قاعدة من المعلومات التي تلزم لمتخذي القرارات حول مدخلات وعمليات ومخرجات أي عملية ، وبالتالي المساهمة في التخطيط وصياغة الأهداف واختيار الأنشطة والأساليب الملائمة وبذلك يعتبر القياس من أهم الركائز التي تساهم في تحقيق الأهداف العامة للمجتمع.

كما تساهم عمليات القياس وبشكل رئيسي في إثارة وتحفيز الأفراد كالطلبة والمعلمين والتربويين الاكاديميين ومتخذي القرارات وراسمي سياسة المجتمع لبذل الجهد المطلوب للوصول إلى الأهداف المرجوة من خلال مساعدتهم في تحديد مواطن العمل والجهود التي تحتاجها وكذلك متطلبات العمل فيها ، مما يعني حثهم على تحسين أساليبهم وإجراءاتهم المستخدمة ، إضافة إلى التحفيز على الجهد والتركيز والتعاون بين الأفراد على اختلاف مواقعهم.

ومن هنا فان القياس يسهم في الوقوف على فاعلية الاجراءات التي تتم ضمن أي مؤسسة والتأكد من مدى فاعليتها من حيث تحديد الإنجازات التي تم تحقيقها

ومدى صحة ودقة وموضوعية إجراءات التحقق ووصف الأوضاع الحالية لها وما هي مواطن القوة والضعف ، وما تتطلبه من إجراءات تطويرية أو تبني سياسات وإجراءات جديدة، من هنا نرى أن هنالك مجالات تطبيقية متعددة ومتباينة لعمليات القياس في أي مجتمع ومنها النظام التربوي مثلا وضمن أي مستوى من مستوياته، وضمن أي مكون من مكوناته.

نشأة وتطور نظرية القياس

يعتبر القياس بإجراءاته وتفصيلاته ضرورة إنسانية وحياتية لا يمكن الاستغناء عنها لأنها بطبيعتها بديهية ومنها ما هو بسيط كمعرفة أسماء الأشياء ومنها المعقد المتعدد الجوانب ككميات الأشياء الصغيرة وخصائصها، وهذا يعني إن هذه الضرورة أي عملية القياس بدأت مع بداية وجود الإنسان على هذه المعمورة ، لكن وبالرغم من ذلك فانه من الصعب تحديد البداية التاريخية لحركة القياس والتقويم ، لكن الأدبيات المتعلقة بنشأة القياس وهنا نقصد القياس النفسي تشير إلى أن تطور هذا العلم مر بست مراحل (Leonard., 1957) إذ أن بدايات القياس النفسي بدأت من خلال الاهتمام بمكونات وخصائص الشخصية الإنسانية وتحديدا بالقدرات وكان أكثر هذه القدرات والتي كانت مدخلا للقياس هي الذكاء، فقد كانت أول هذه المراحل هي مرحلة زمن الفلاسفة أمثال سقراط وافلاطون وقد أطلق عليها المرحلة التأسيسية أو الأولية " Historical Foundations " وقد امتدت هذه المرحلة من (٣٤٨ قبل الميلاد) وحتى العام (١٨٣٣) بدءا بالفيلسوف أفلاطون وانتهاء بالفيلسوف جسبرد آيتارد (Gaspard,Itard,1775) ومن علماء هذه المرحلة باسكال(Pascal,1623) وكانت (Kant,1724) حيث أوردت المراجع اهتمامهم بدراسة شخصية الإنسان وتحديدا في الذكاء أو القدرة العقلية.

أما المرحلة الثانية والتي أطلق عليها المرحلة التأسيسية الحديثة Modern Foundations Period على اعتبار أن المرحلة الأولى كانت في العصر قبل الميلاد أيام اليونانيين فقد بدأت هذه المرحلة في العام (١٧٠٤) بدءاً بالفيلسوف والمنظر السياسي جان لوك (Locke,1632) الذي أشار إلى أهمية الخبرة في تنمية القدرة

العقلية وانتهاء بالطبيب الفرنسي شاركوت (Charcot١٨٢٥). أما المرحلة الثالثة والتي أطلق عليها مرحلة مدارس علم النفس The Great Schools Period حيث برز فيها نشاط العلماء الأوروبيون إضافة إلى دراسة العلماء الأمريكان لعلم النفس خارج أمريكا والعودة لتطبيق ما تعلموه والتوسع فيه في أمريكا حيث بدأت هذه الاستخدامات في القوات المسلحة ومن علماء هذه المرحلة فونت (Wundt1,1832) وابنجهاوس (Ebbmghaus,1850) إضافة إلى كاتل (Cattell,1860) أما المرحلة الرابعة فقد بدأت مع ثورونديك (Thorndike,1847) وبينيه(Binet,1857) وانتهت مع كوكس (Cox,1890) وبياجيه (Piaget,1896) وأطلق عليها مرحلة انتشار تأثير مدارس علم النفس The Great Schools' Influence Period حيث انقسم العلماء حسب وجهات النظر المختلفة وبدأت هذه الآراء تمتد إلى نطاق أوسع من حيث المؤيدين والمعارضين، واهم منجزات في هذه المرحلة تطوير اختبارات ألفا وبيتا في الجيش الأمريكي تحت إشراف روبرت ايركس (Yerkes,Robert,1876) حيث كان المجال خصبا لاكتساب الخبرة في بناء الاختبارات، إضافة إلى اقتراح التصاميم التجريبية للبحث في مجال القياس النفسي وتصميم الاختبارات

أما المرحلة الخامسة والتي عرفت بمرحلة الاستكشافات العصرية contemporary Explorations حيث ظهر كل من بيرت وزيج وكاتل وجلفورد وثيرستون وبينيه وفيرنون (Burt,1883&Psyche,Cattell,1893&Gilford, 1897&Thursone, 1897&Vernone,1905) وفي هذه المرحلة تمكن بعض العلماء من تطوير بعض الأساليب الإحصائية مثل سبيرمان وثروندايك (Sperman,1863 & Thorndike,1847) وغيرهم وذلك لتحليل البيانات ومعايرة نتائج الاختبارات تمهيدا لتجسيد صدق وثبات الاختبارات . أما المرحلة السادسة وهي المرحلة الحالية والتي وصل تطور القياس إلى أوجه حيث تطورت البرامج والأساليب الإحصائية وكذلك برزت النظرية الحديثة (Item Response Theory) وقد تخللت هذه المراحل أنشطة متعددة للعلماء في مجال القياس النفسي حيث كانت البدايات مع القدرة العقلية باعتبارها من مكونات شخصية الفرد

إضافة إلى الاهتمام بالتصنيف والترتيب ، ثم توالت هذه الأنشطة تأخذ طابع الاختصاص فكانت البدايات تهتم بتصنيف الأشخاص وفرزهم إذ كان لا بد من توفر أدوات لتحقيق ذلك ومن هنا بدأ تصميم الاختبارات واستنتاج الإجراءات الإحصائية اللازمة لتحليل البيانات المتحصلة نتيجة للاختبارات حيث أصبحت الأساليب الإحصائية أكثر ضرورة خاصة مع تزايد أعداد وتنوع الأفراد الذين تطبق عليهم الاختبارات

وهنا يمكن القول أن ولادة علم القياس بدأت في الحقل الإنساني ، فقد استخدمت الاختبارات منذ قديم الزمان لتساعد في التعرف على خصائص الأفراد ولا سيما الطلبة لتحديد الذين يستطيعون أن ينتقلوا من صف الصف التالي أو من مرحلة دراسية إلى أخرى في السلم التعليمي وكذلك لتحديد الأفراد الذين يحتاجون لبرامج تعليمية معينة .

وقد مارست المجتمعات نشاطات القياس في مختلف العصور فقد بقي قياس التحصيل غير منظم في أهدافه وإجراءاته عند الصينيين حتى سجلت أول حادثة للقياس الرسمي المنظم عند الصينيين القدامى في فترة ما قبل الميلاد حين لجئوا إلى الامتحانات العامة لاختيار موظفي الحكومة . وكان يحق لكل من يجد في نفسه القدرة أن يتقدم لامتحانات الدرجة الأولى فإذا اجتازها بنجاح أصبح مؤهلا لتسلم مركزا حكوميا في مجتمعه المحلي الذي يسكن فيه . فإذا أراد إشغال وظيفة أعلى فإنه يمكن أن يتقدم لامتحانات الدرجة الثانية فالثالثة وهكذا وقد كانت امتحانات الدرجة الأولى تقام مرة كل ٣ سنوات ومدتها ٣ ساعات ، وكان يطلب من الطالب أن ينشئ رسائل على غرار كتابات كونفوشيوس وأما امتحانات الدرجة الثانية فتقام مرة كل ٣ سنوات أيضا ومدتها ٣ أيام وموضوعاتها أشمل من امتحانات الدرجة الأولى .أما امتحانات الدرجة الثالثة فكانت موضوعاتها أشمل من امتحانات الدرجة الثانية وتستمر 13 يوما وتطبق في بكين بقاعة (هان لين)، أي غابة الأقلام ، وكان الناجح يتبوأ أعلى المناصب العليا والمهمة في الدولة

أما في مصر القديمة فقد اتجهت الاختبارات اتجاهين الأول يتمثل في الامتحانات العملية والثاني في الامتحانات الكتابية التي كان يقوم بها الكهنة في المعابد ، وفي المدارس والمعاهد العليا التابعة للحكومة . وقد اتخذت الاختبارات صنفين رئيسين الأول يشبه كما يعرف اليوم بالتلمذة المهنية غير الرسمية والتربية الدينية

أما القياس لدى الأثينيين القدامى فقد مارسه المعلمون أمثال سقراط وأفلاطون وأرسطو وبركليت حيث كانوا يعلمون الشباب الأثيني ويقيمون معرفتهم بأساليب لفظية تعتمد على الحوار ، فقد اعتقد سقراط على سبيل المثال ان العلم تذكر والجهل نسيان ، وأحسن وسيلة للتذكر إنما هي الجدل (الحوار)، وقد شبه سقراط طريقته هذه بعملية الولادة وشبه نفسه بوالدته القابلة التي كانت تخرج الصغار من بطون النساء بينما كان هو يخرج الأفكار من عقول الرجال ويشير الجزء الخاص بإعداد الشخص المنتج والحارس والحاكم الفيلسوف في جمهورية أفلاطون إلى استخدام أساليب القياس والاختبارات لتوزيع أفراد المجتمع على الفئات الثلاث التي يتكون منها أفراد الجمهورية.

أما الرومان فقد اهتموا بالامتحانات الشفوية العملية أكثر من غيرها لأنهم اهتموا بإنتاج الخطباء المفوهين وعادة ما كان الطالب يلقي خطبته علنا وعلى قارعة الطريق حيث يجتمع نفر من الناس بالقرب من المدرسة التي كانت تطل على الطريق مباشرة ومن خلال براعة الخطبة وتأثيرها كان يحكم على الخطيب، حيث تعتبر الخطبة إلى ما يشبه المقابلة من طرف واحد.

القياس لدى المسلمين

تكمن أهمية القياس والتقويم في الفكر الإسلامي كونه الأداة الرئيسة لصنع القرار السليم في جميع مناحي الحياة ، حيث أن اتخاذ القرار والحاجة إليه أمر ملازمٌ لحياة الأفراد والمجتمعات، بدءاً من الممارسات الحيوية اليومية مروراً بالخيارات المؤسسية وانتهاءً بالقرارات المصيرية.

إن المتأمل في مفهوم الحياة وصياغة فلسفة الوجود في الدين الإسلامي يلاحظ وبشكل واضح الهدف منه محاسبة النفس ومراجعتها بشكل ذاتي وانتهاء بإتقان العمل

وجوده ، وقد خلق الله الإنسان لحكمة بالغة هي الاختبار﴿ الَّذِي خَلَقَ الْمَوْتَ وَالْحَيَاةَ لِيَبْلُوَكُمْ أَيُّكُمْ أَحْسَنُ عَمَلًا وَهُوَ الْعَزِيزُ الْغَفُورُ ﴾(1) ، كما أوضح الله عز وجل أن الهدف من وجود الإنسان هو النجاح في الاختبار او الغاية من الحياة الدنيا المتمثل في تحقيق العبادة، ﴿ وَمَا خَلَقْتُ الْجِنَّ وَالْإِنسَ إِلَّا لِيَعْبُدُونِ ﴾(2) هذا من جهة أصل الوجود.ومن جهة التعامل مع الأفراد والأمم والأفكار فقد جاء الإسلام بمنهج علمي رصين يقوم على أساس العدل والإنصاف حتى مع الأعداء، فلقد تواترت نصوص القرآن والسنة، على أهمية العدل والموضوعية والبعد عن الذاتية في إصدار الأحكام تجاه الآخرين. ومن ملامح هذه الموضوعية والبعد عن الذاتية التي أمر بها الإسلام ما يلي.

- أهمية التثبت﴿ يَا أَيُّهَا الَّذِينَ آمَنُوا إِن جَاءَكُمْ فَاسِقٌ بِنَبَإٍ فَتَبَيَّنُوا أَن تُصِيبُوا قَوْمًا بِجَهَالَةٍ فَتُصْبِحُوا عَلَىٰ مَا فَعَلْتُمْ نَادِمِينَ ﴾(3). وهذا يعني صدق المعلومة باعتبار انه سيتخذ حكما وهو بمثابة القرار حيث سيعتمد على صدق المعلومة.
- البُعد عن الظن، ذلك أن الظن لا يفيد اليقين ومن الخطأ بل ومن الظلم بناء الأحكام او اتخاذ القرارات بناء على مصادر ظنية غير قطعية ﴿ يَا أَيُّهَا الَّذِينَ آمَنُوا اجْتَنِبُوا كَثِيرًا مِّنَ الظَّنِّ إِنَّ بَعْضَ الظَّنِّ إِثْمٌ ﴾(4). وهذا يعني البحث والتأكد من المعلومات وعدم تركها للظن والعشوائية وهو منهجية علم القياس من حيث جمع المعلومات بشكل فعلي ، إضافة إلى أن عدم الظن يعني التحري الذي يوازي وجود أداة لهذا التحري وهذا يندرج ضمن أداة القياس.
- الأمر بالقسط والعدل حتى مع الأعداء، قال تعالى: ﴿ يَا أَيُّهَا الَّذِينَ آمَنُوا كُونُوا قَوَّامِينَ لِلَّهِ شُهَدَاءَ بِالْقِسْطِ وَلَا يَجْرِمَنَّكُمْ شَنَآنُ قَوْمٍ عَلَىٰ أَلَّا تَعْدِلُوا اعْدِلُوا هُوَ أَقْرَبُ لِلتَّقْوَىٰ ﴾(5) حيث ان اتخاذ القرار السليم هو احد أشكال العدل

(1) سورة الملك: الآية ٢

(2) سورة الذاريات: الآية ٥٦.

(3) سورة الحجرات: الآية ٦

(4) سورة الحجرات: الآية ١٢

(5) سورة المائدة. الآية ٨.

خاصة إذا ما تعلق بحقوق الأفراد ، ولن يتحقق العدل دون وجود أداة قياس صادقة وموضوعية.

وهذه إشارات فيما يتعلق بالحكم على الأشخاص وتقويم الأفكار وصنع القرار، أما فيما يتعلق بالقياس المادي الطبيعي، فإن المتتبع لنصوص القرآن الكريم والسنة النبوية يجد عشرات النصوص في ذلك مثل إنزال الرزق بقدر معلوم، ﴿ وَإِن مِّن شَيْءٍ إِلَّا عِندَنَا خَزَائِنُهُ وَمَا نُنَزِّلُهُ إِلَّا بِقَدَرٍ مَّعْلُومٍ ﴾ [1] أي بقدر يمكن أن يفهم مقداره وكمه، ومن ذلك تقدير منازل القمر وعدد الشهور وأنصبة الزكاة والفرائض وأحكام المواريث وسن الزواج وعدة الطلاق بأنواعه ومدة الحمل وغيرها

كما إن علم الحديث وعلم الجرح والتعديل يعتبران أنموذجاً لتطبيقات علم القياس ، فمن خلالهما يتم تحويل الصفات (ألفاظ الجرح والتعديل) إلى مراتب وأرقام، كما مارس المحدثون النقد الخارجي (المحكات) لدراسة سند الحديث، ومارسوا النقد الداخلي (المعايير) لدراسة متن الحديث.

أما بالنسبة للمسلمين ومن الناحية الإجرائية فقد ظل تقييم التحصيل يعتمد بدرجة رئيسة على الاختبارات الشفوية كالتسميع إلى جانب اختبارات الكتابة والخط العربي. ففي المساجد أتيحت للطلاب دائما فرصة الاستماع إلى الأساتذة ا لزائرين من كل أنحاء العالم الإسلامي آنذاك ولم تكن هناك شهادات كما هو معروف الآن ، وإنما كان الطلاب يدرسون كتبا معينة على يد أساتذة مختصين فإذا ما استطاع الطالب استيعاب الكتاب الذي يدرسه إياه شيخه فعندئذ يكتب له شهادة على الورقة الأولى أو الأخيرة من الكتاب تسمى الإجازة يبين فيها أن الطالب المذكور قد أتم قراءة الكتاب وأجيز له تدريسه

القياس في العصور الحديثة .

إنشغل علماء الطبيعة في أواخر القرن التاسع عشر في تصنيف النباتات والحيوانات، مما دعا علماء ذلك العصر في بريطانيا لأن يوجهوا انتباههم إلى محاولة

---

(1) سورة الحجر: الآية ٢١

قياس الفروق بين الناس كذلك. حيث كان العلماء يحاولون دراسة خصائص الشخصية فقد أنشأ وليام فونت (Wundt,1832) في النصف الثاني من القرن التاسع عشر أنشأ أول معمل لعلم النفس التجريبي في مدينة ليبزج بألمانيا وكان ذلك في عام (١٨٧٩) ، أما أول من حاول قياس القدرات البشرية علمياً هو فرانسيس غالتون(Gallton,1822). ويعود له الفضل في وضع حركة القياس في مجراها الصحيح حيث جمع كما ضخما من البيانات عن الفروق الفردية ، وفي (١٨٨٢) أنشأ معملا لعلم النفس على غرار معلمه فونت اقتصرت أعماله فيه على الأفراد الموهوبين اذ كان متأثرا بأفكار داروين حول الوراثة والتي جمعها في كتابه المعروف 'عبقري بالوراثة' لذلك اهتم بدراسة الوراثة عند الإنسان واكتشف درجة تشابه الذرية في الخصائص الشخصية كالتوائم والإخوة الأخوات وأبناء العمومة

وكان جالتون أول من طبق منهاج الاستبيان والمقياس المتدرج ومنهج التداعي الحر ومقاييس التمييز الحسي، وقد استخدم مقاييس التمييز الحسي لقياس الذكاء ، فهو يرى أن المعلومات التي تصل إلينا عن العالم من حولنا هي التي تمر عبر حواسنا . ويكبر حجم تمايز هذه الحواس تتسع القاعدة التي نستقي من خلالها معلوماتنا والتي نبني عليها أحكامنا وذكاءنا. ونتيجة لذلك حاول غالتون تقييم الآلاف من الناس باستخدام مجموعة من الاختبارات البدائية كان معظمها لتقييم رد الفعل وقدرات التنسيق والمهارات الحركية وكانت نتيجة هذه التجارب الاهتمام بحركة القياس وعلى الأخص اهتمام بعض الدول والتي كانت تسودها بعض التوجهات العنصرية. وقد كان هدف غالتون فهم ماهية العبقرية وطبيعتها وكيف يمكن تحسينها من خلال تربية نوعية منتقاة من الأفراد وكان ذلك سنة (١٨٦٩).حيث أن اهتمامه بالوراثة قاده إلى التساؤل عن كيفية تحسين النسل الأمر الذي دفع بالفكرة التي مفادها أن تحسين النسل يمكن أن يتم من خلال الاهتمام بانتقاء الأب والأم أي أن الوراثة لها دور في تشكيل عبقرية الفرد ، كما أشار إلى أن الخصائص البشرية تتوزع بشكل طبيعي وهذه إشارة إلى التوزيع الطبيعي.

وتبعا لذلك، وجهت الانتقادات – التي تعرض لها هذا الاستخدام لهذه الاختبارات – علماء النفس إلى الابتعاد عن التقليد الذي كان يصنف علم النفس كفرع من فروع علم الفلسفة والاهتمام به كأحد العلوم الطبيعية. مع أخذ هذا الهدف في حين الاعتبار، كانت أهم أهداف هذا العلم هو التخلي عن تقليد استكشاف أي ظاهرة نفسية عن طريق التأمل الذاتي أو الفلسفي والاستعاضة عنها بتطبيق مبادئ البحث العلمي ومن أجل تحقيق ذلك اضطر العلماء إلى محاولة الوصول إلى قياس الظواهر النفسية وهو مجال القياس النفسي (قياس الخصائص النفسية)

وفي هذه الفترة بدا الاهتمام باختبارات التحصيل انطلاقا من الاهتمام بقياس السمات العقلية حيث كان العالم الأمريكي جوزيف رايس(Josef, Rice) الرائد الأول لاختبارات التحصيل الموضوعية وذلك باختباره للهجاء الذي ظهر عام (١٨٨٩)إذ كانت الاختبارات السائدة آنذاك تقتصر على المشافهة .ففي عام (١٨٩٧) طلب من ابنجهاوس (Ebbinghaus,1850) وهو من مؤسسي علم النفس التجريبي أن يدرس حالات التعب عند بعض الأطفال ولهذا الغرض قام بتصميم ثلاثة اختبارات لقياس كل من الذاكرة والحساب وإكمال الجمل حيث أشار إلى ارتباط نتائج اختبار تكملة الجمل بالنتائج المدرسية ، وفي نفس الوقت قام ديسانتس (Desanctis) بنشر ستة اختبارات خاصة بضعاف العقول تتعلق بذاكرة الألوان وتتبع العمليات العقلية والتفكير ، وبقيت تستخدم حتى نشر كل من بينيه وسيمون (Binet,1857) وهنري (Henri) اختبارهما المعروف باختبار بينيه

وفي بدايات القرن التاسع عشر(١٨٠٠) أثار اسكيرول (Esquirol,1772) فكرة الفروق بين القصور العقلي والعته وقدم إرشادات إلى كيفية التعامل مع الأشخاص في المصحات النفسية حيث كان يعمل في احد السجون واستخدم القدرة اللغوية في التميز بين الأشخاص ذوي التخلف العقلي من خلال مقاييس الجمجمة وملامح الوجه، حيث أشار إلى أن بعض مستويات التخلف العقلي يمكن تطويره عن طريق التدريب.

وفي نفس الفترة تقريبا برزت اهتمامات مارتن كاركوت J,Martin (Charcot,1825) بتشخيص أسباب الهستيريا والنوبات النفسية من الناحية العصبية .إذ ربط بين النواحي النفسية والعمليات الفسيولوجية للجسم، ولذلك يعتبره البعض المؤسس لنظريات التحليل النفسي من خلال استراتيجيات التعامل مع المصابين بحالات الهستيريا في مستشفى الطب النفسي الذي عمل فيه

ومن الذين ساهموا في تطور القياس العقلي أيضا فونت (Wundt,1832) الذي يعتبر من أهم رواد علم النفس التجريبي ومن مؤسسي علم النفس كما يعتبر أول من أسس مختبرا تجريبيا في علم النفس وتم اعتبار مختبره نموذجا لمختبرات تم تأسيسها فيما بعد ، إضافة إلى اهتمامه بقياس القدرة على التمييز الحسي كما اهتم بالعلاقة بين الإحساس والقوة البدنية ، كما اقترح فونت أسلوبا لتحديد مدى حساسية الفرد لبعض المثيرات والتي استفاد منه بينيه فيما بعد عند تحديد العمر الزمني والعقلي للمفحوص على الاختبار الذي طوره لقياس القدرة العقلية.

كما اهتم عالم النفس الألماني هيرمان ابنجهاوس (Hermann Ebbinghaus,1850) بدراسة الذاكرة الإنسانية والعمليات العقلية العليا وأشار إلى انه يمكن دراستها وتحليلها بشكل علمي، وهذا ما ساهم في نقل مفهوم الذاكرة من المجال الفلسفي إلى المجال العلمي التطبيقي وقد استخدم المقاطع الرمزية في تنمية الذاكرة إذ طور (٢٣٠٠) رمز لتنمية الذاكرة ومعرفة الزمن اللازم للتذكر في حالة النسيان، وقد توصل ابنجهاوس نتيجة لتجاربه على الذاكرة إلى أن الزمن اللازم للتذكر أو للتعلم في حال النسيان اقل والتعلم اللاحق أسهل من التعلم للمرة الأولى ، كما أشار إلى أمر مهم وعنوان التعلم الناجح يقتضي إعطاء جرعات أو مهمات تعليمية(دروس) على فترات متقطعة بدل أن تكون في فترة متواصلة .وهذا بحد ذاته كان تمهيدا لطرح مفهوم الفروق الفردية خاصة من خلال اهتمامه بدراسة زمن رد الفعل والإدراك الحسي .

وفي هذه المرحلة كان عالم النفس الأمريكي جيمس كاتل(Cattel,j.1860) قد درس علم النفس في أوروبا وعاد إلى أمريكا وحاول تطبيق ما عرفه في الولايات

المتحدة ، ويعتبر كاتل من رواد علم النفس الأمريكي وقد قدم لعلم النفس ما لم يقدمه احد على امتداد القرن الثامن عشر ، حيث كان يشغله هدف مهم بالنسبة له وهو تجسيد فكرة أن علم النفس هو علم لا يختلف عن العلوم البحتة ، وقد عمل على تحقيق هذا الهدف مستفيدا من أهمية علم النفس بالنسبة لبناء المناهج على مستوى المدارس والجامعات ، كما عمل على نشر فكرة تأسيس المجلات المتخصصة بعلم النفس على غرار تلك المجلات التي كانت منتشرة لكل من الرياضيات والبيولوجي والفيزياء ، وقد أسس الرابطة الأمريكية لعلم النفس ورابطة أساتذة الجامعات والرابطة الأمريكية لتطوير العلوم كما اعتقد كاتل بدور الطرق الكمية والإحصائية في تطوير علم النفس كباقي فروع العلوم البحتة ، إضافة إلى دور التجربة في بيان أهمية علم النفس وبسبب تشكيك الكثيرين آنذاك بعلم النفس وإمكانية القياس فيه استمر كاتل بطرح فكرة علم النفس التجريبي وإمكانية قياس السمات النفسية ولا سيما الخصائص الجسمية والنفسية، حيث كانت وجهة نظرة هذه والتي تتمحور حول قياس واختبار الخصائص البشرية ناتجة عن تأثره بمعلمه غالتون الذي عمل في مختبره في ليبزخ بألمانيا وفي عام(١٨٩٠) استخدم كاتل اصطلاح الاختبار العقلي لأول مرة من خلال مقالته التي نشر فيها عن أكثر من خمسين اختبارا كان يطبقها على الطلبة في الكليات لتحديد مستوياتهم العقلية، وأشار إلى أن الوظائف العقلية يمكن أن تقاس عن طريق اختبارات التمييز الحسي ، وسرعة وزمن رد الفعل . إلى جانب ذلك ركز على مفهوم الفروق الفردية وكان له الفضل في استخدام الاختبارات لقياس الاستعدادات العقلية وقد اختلف كاتل مع أستاذه فونت في مسألة الفروق الفردية حيث كان يرى أنها حقيقة واقعية وانه يمكن قياسها ، في حين كان فونت قد اعتبرها نتيجة لأخطاء القياس والتغير في الطبيعة ، إلى جانب ذلك كان جوزيف جاسترو (Joseph.Jastrow,1893) يطبق اختبارا بسيطا على زوار معرض شيكاغو وهم من جنسيات مختلفة وكان يتضمن فقرات حسية إدراكية وقد أشار إلى مفهوم المعايير وبناء اختبارات للمهارات الحركية.

كما اهتم عالم النفس الألماني كريبلن (Kraepelin,1947) بقياس العمليات العقلية المعقدة، ومن هنا كان اهتمامه منصبا على بطاريات الاختبارات التي تفحص السمات العقلية كالتذكر والتشتت والتعب حيث كان مهتما في علم النفس المرضي، وفي هذا المجال قام اورن (Oehrn,1895) بتصميم مجموعة من الاختبارات لقياس كل من الإدراك والتذكر والوظائف الحركية لدراسة العلاقة بين الوظائف النفسية ، ومع مرور الوقت، تغير مفهوم القياس النفسي ليكون مخصص لوصف النماذج الإحصائية التي تدعم تصميم واستعمال الاختبارات التي تجرى بالورقة والقلم، بدلاً من الاختبارات التي تستخدم لوصف القياس النفسي في أوسع معانيها (مثل التقييم المهاري و الذاكرة والانتباه ..الخ). وقد وجد لغاية الآن نموذجان إحصائيان أساسيان يدعمان تصميم الاختبارات أولهما النظرية الكلاسيكية للاختبارات*(CTT) (تسمى أيضاً نظرية النتائج الفعلية) والتي تطورت من أعمال سبيرمان (١٩٠٤) ، والنظرية الثانية هي نظرية الاستجابة على الفقرة (IRT) والتي تطورت من أعمال (Rasch,١٩٦٠) والتي تسمى بنظرية السمات الكامنة أيضاً، أو نظرية النماذج اللوغريتمية.

اما أول بادرة مقننة لتقييم التحصيل فقد كانت على يد جورج فيشر الذي ألف كتاب "الميزان" عام (١٨٤٨) . احتوى على مقياس للكتابة اليدوية يمكن به تبويب مهارات الطلاب الكتابية ، واحتوى كذلك على مجموعات من الأسئلة في علوم الرياضيات والملاحة والقواعد والإنشاء واللغة الفرنسية والتاريخ والرسم .

في نهاية القرن التاسع عشر برزت اهتمامات ادوارد ثورندايك (Thorndike,1874) بدراسة سلوك الحيوان حيث عكف على ملاحظة سلوك القطة اللغز ('cats in a puzzle box') وحاول الاستفادة من سلوك الحيوان لفهم سلوك البشر وكان أن خرج بقانونه المعروف قانون الأثر والخاص بتعزيز السلوك الايجابي والعقاب على السلوك الخاطئ كما استخدم ثورندايك مع طلابه تحديد الهدف من قياس الذكاء في وقت مبكر كما وضع أساليب لقياس عدد من القدرات والإنجازات.

---

* للمزيد حول وصف للنظرية الكلاسيكية يمكن الرجوع الى كلاين (Cline,1986) .

وفي عام ١٩٢٠ وضع اختبار الذكاء المكون من فقرات التكميل والحساب ، والمفردات، واختبار الاتجاهات ، والمعروفة باسم (CAVD). وكان الهدف من ذلك قياس الذكاء على مقياس مطلق وتحديد العوامل الكامنة وراء الذكاء وأصبح فيما بعد لبنة لقياس الذكاء كما أشار ثورنديك إلى ثلاثة أنواع من الذكاء هي الذكاء المجرد والذكاء الميكانيكي(الجسمي) والذكاء الاجتماعي.

واستمر اهتمام علماء القياس النفسي بقياس الذكاء حيث نشر بينيه (Binet,1857) مقالا انتقد فيه الاختبارات التي كانت سائدة وتستخدم لقياس الذكاء وكان ذلك عام (١٨٩٥) واقترح قائمة متنوعة من الاختبارات التي تشمل على وظائف عقلية مختلفة كالتذكر والتخيل والانتباه والفهم بالاشتراك مع سيمون حيث كان بينيه الطبيب الفرنسي الذي كان يبحث في الأعصاب فدرس القانون وتفرغ ليتعلم علم النفس بشكل ذاتي ونتيجة لمطالعاته جذبت انتباهه أفكار جون ستورث (Stuart,1806) والتي تلخصت بأن الذكاء الإنساني يمكن فهمه وتفسيره من خلال قوانين الارتباط ، كما اهتم بدراسة الطفل وكان يدير مختبر علم النفس التجريبي في جامعة السوربون، وهو أول من دعا وانشأ مجلة علم النفس الكبرى في العام (١٨٩١)، ونتيجة لنشاطه طلب وزير التربية والتعليم الفرنسي في عام (١٩٠٥) تشكيل لجنة لعزل الطلبة المتأخرين دراسيا في المدارس الفرنسية وقام بالمهمة بالاشتراك مع سيمون وكان الاختبار الذي استخدموه اختبارا فرديا مكونا من (٣٠) سؤالا مرتبة بشكل تصاعدي من حيث الصعوبة واشتهر وبقي يعرف باسمة حتى أيامنا الحالية حيث أجريت عليه فيما بعد عدة تعديلات من قبل لويس تيرمان(Terman,1877) الذي كان يعمل في جامعة ستانفورد حيث عرفت النسخة المعدلة من الاختبار فيما بعد باختبار (ستانفورد- بينيه) .أوكلت إليه الحكومة الفرنسية بإيجاد آلية لتحديد الأطفال من ذوي التخلف العقلي لإيجاد منهاج بديل يتعلمون من خلاله وبالفعل قام بينيه بدراسة ذلك وساعده بذلك تجربته مع ابنتيه وشاركه في ذلك سيمون (Simon,1873) حيث انتهت الجهود إلى تطوير مقياس القدرة العقلية الشهير والذي حمل أسمه مقياس بينية للقدرة العقلية وكان ذلك في عام١٩٠٤، ومن ثم

جاءت مساهمات عالم الرياضيات البريطاني كارل بيرسون (Pearson,1857) الذي اهتم بأفكار فرانسيس جالتون حول الارتباط بين الصفات البيولوجية عبر الأجيال حيث انصب اهتمامه على إيجاد أساليب إحصائية تفسر ذلك وفي بداية العام (١٩٠٠) توصل إلى العديد من الأساليب منها معامل الارتباط ومعامل ارتباط بوينت بايسيريال والانحراف المعياري والارتباط الجزئي والارتباط المتعدد ومعامل التغير إضافة إلى مفهوم التفلطح في التوزيع الطبيعي، حيث بدأ استخدامها في تحليل البيانات

بعدها ازداد اهتمام علماء النفس بالأساليب الإحصائية ومفاهيم القياس التي تفسر نتائج قياس سمة الذكاء وقد بدأت مفاهيم القياس (Measurement) مع عالم النفس الانجليزي تشارلز سبيرمان (Sperman,1945) الذي درس الهندسة في سن متأخرة نسبيا وقد اعتبره عالم النفس الأمريكي جنسن (Jensen,1923) الأب الروحي لنظرية القياس التقليدية (Classical Test Theory) وذلك من خلال مساهماته في استنتاج العامل العام كمكون من مكونات القدرة العقلية (الذكاء) وذلك من خلال الأسلوب الإحصائي الذي اسماه التحليل العاملي (factor analysis) والذي يستند إلى تحليل الارتباط بين مكونات الذكاء، ويبقى هذا الأسلوب من أهم الأساليب الإحصائية حتى بدايات القرن الواحد والعشرون، إضافة إلى إشارته إلى مفهوم معامل الارتباط المصحح وذلك عندما تتجمع البيانات ضمن مدى معين من الدرجات، وهذا ما ستنتجه عندما حلل بيانات كاتل (Cattel) وصحح معاملات الارتباط بين بعض المتغيرات، وقال حينها أن ضيق المدى للدرجات يقلل من قيمة معاملات الارتباط البينية بين تلك المتغيرات وقد كان تعريف وتفسير سبيرمان للذكاء أول تعريف مقنع وفيه من المنطق ومدعم بالأساليب الإحصائية، وهو صاحب نظرية العامل العام في الذكاء والذي عارضه فيما بعد أقطاب نظرية الذكاء المتعدد

في عام ١٩٠٨ قام عالم النفس الأمريكي جودارد (Goddard,1866) بإدخال اختبار سيمون – بينيه إلى الولايات المتحدة وقام بترجمته إلى الانجليزية وطبقه على عينة كبيرة جدا وصلت إلى (٢٢٠٠٠) طفل وقام باستنتاج معايير له على البيئة الأمريكية وكان ذلك في جامعة ستانفورد واخذ الاختبار اسم مطوره الأصيل بينيه واسم الجامعة

التي طور فيها وأصبح اسم الطبعة الثانية منه باسم مقياس ستانفورد – بينية (Stanford-Binet Scale) وخرج الاختبار بصورته الجديدة عام(١٩١٨) ولذلك يعتبر جودارد اختبارات الذكاء الأمريكية ومن الجدير بالذكر أن جودا رد كان يعتقد بأن الكروموسومات هي المسؤولة عن تحديد الذكاء عند الإنسان وهي تتأثر بالأحداث التي تدمر بعض الكر وموسومات وبالتالي تأثر قدرات الفرد العقلية.ولذلك دعا جودا رد إلى ضرورة تحسين النسل في أمريكا وكذلك دعا إلى وضع قانون يضبط الهجرة إلى أمريكا لاعتقاده بان نسبة كبيرة من المهاجرين إلى أمريكا مصابون بالتخلف العقلي .

في عام ١٩٠٦ قام تيرمان (Terman,1877) وهو ناشط في علم النفس المقارن بتكييف اختبار ستانفورد- بينيه ليكون اختبارا جماعيا للبيئة الأمريكية وطبقه على مجموعة من الأطفال وكان يهدف إلى معرفة مدى نجاح الأطفال الذين لديهم ذكاء عالي في وقت لاحق ، وقد أجرى دراسته لمرحلة الدكتوراه في هذا المجال وكانت تحت عنوان العبقرية والغباء حيث عمل من خلالها على تصميم اختبارات لقياس ما سماه بالعمليات العقلية السبع آنذاك ، وقد توصل إلى أن هناك عوامل أخرى غير الذكاء تساهم في النجاح المستقبلي للفرد ، وقد وافق ستيرن(Stern) في مفهوم نسبة الذكاء التي سميت فيما بعد بنسبة الذكاء الانحرافية (IQ).

وفي نفس الفترة اضاف فيجوتسكي (Vygotsky,1896) وهو عالم نفس روسي تفسيرا آخر للذكاء واعتبره على انه مجموعة من الأنشطة العملياتية وتلعب البيئة الاجتماعية دورا مهما في تطويره ولم يؤيد النظرة للذكاء على انه مجموعة من البنى العقلية الكامنة ، كما قال بان ذكاء الإنسان يمكنه من التكيف مع البيئة التي يعيش فيها وكذلك يمكنه من التأثير فيها، بمعنى أن الذكاء أنشطة عملياتية أكثر منه قدرة.

وفي نفس الفترة تقريبا طرح وليام ستيرن (Stern,1871) مفهوم علم نفس الشخصية، لاعتقاده بان القدرة العقلية للفرد تتأثر بالفروق الفردية وقد عرف الذكاء بأنه قدرة الفرد على التفكير في متطلبات محددة وكذلك القدرة على التكيف مع متطلبات الحياة، واهتم بكيفية حساب ما سماه بحصيلة الذكاء (Intelligence Quotient) وذلك من خلال النسبة بين العمر العقلي إلى العمر الزمني.

وقد حاول البعض إيجاد طرقا لقياس الذكاء غير الصورة التي بدت مألوفة وهي الاختبارات الكتابية والشفوية التي شاعت في تلك الفترة ، فقد حاولت عالمة النفس الأمريكية فلورنس جودنفا (Goodenough,1886) أن تقيس الذكاء من خلال الرسوم وكطريقة بديلة للاختبارات اللفظية وفعلا اقترحت اختبارا بقي يستخدم لفترة ليست بالقليلة وسمي باختبار رسم الرجل الذي كان يستخدم كاختبار قبل مرحلة الدراسة وذلك للمرحلة العمرية (٢ - ١٣) سنة وكان أول تطبيق له في مدرسة مينسوتا وقد عرف أيضا باختبار مينسوتا إضافة إلى انه تم تطويره فيما بعد من قبل هاريس (Harris) عام (١٩٤٠) وتم إيجاد صورة أخرى وتضمن إضافة جديدة حملت اسم باختبار رسم المرأة وعرفت النسخة المنقحة باسم (Goodenough-Harris drawing test). كما اهتمت جودا بنسبة الذكاء وأكدت أنها مختلفة من طفل إلى آخر وقد استخدمت في مقارناتها النسب المئوية حيث كانت ترى أنها أسهل من حيث الفهم والمقارنة بين الأفراد.

وفي العقد الخامس من القرن العشرين طرح جان بياجيه (Piaget,1896) المفهوم أو التفسير البيولوجي للذكاء، وعرف الذكاء على انه تكيف الفرد مع البيئة التي يتواجد فيها ، حيث اعتبر أن المعلومات المعروفة للفرد (الخبرة السابقة) يمكن تعديلها بما يتفق والمعلومات الجديدة وهو بذلك يشير إلى فكرة التعديل الداخلي أو الخارجي (ssimilation & Assimilation) حيث كان مهتما وبشكل مبكر في حياة الحيوانات وعمل في متحف العلوم الطبيعية مع بول غودبه وعمل في التحليل النفسي لفترة أربع سنوات كما عمل مع سيمون لتنقيح النسخة الفرنسية لاختبار بينيه.

ومع ازدياد اهتمام العلماء بفهم وتفسير الذكاء ومع تعدد الآراء بدأ الاهتمام يتوجه إلى دراسة الأشخاص من ذوي الذكاء العالي او ما سمي بالعبقرية حيث حاولت عالمة النفس الأمريكية كوكس (Cox,1890) بدراسة نسب الذكاء للأفراد التي كانت تعتقد أنهم مبدعون أو عباقرة وذلك من خلال تحليل السير الذاتية لهم ، حيث حاولت تقدير نسب الذكاء لحوالي (٣٠٠) شخص افترضت أنهم عباقرة ، إضافة إلى تحليل (٦٧) خاصية ل(١٠٠) شخص وخلصت إلى أن نسبة الذكاء تتأثر بالدافعية

والصحة الجسمية والنفسية ، ويبدو أن عملها مع تيرمان في الدراسة الطولية على الأطفال المبدعين دفعها إلى الاهتمام بدراسة العبقرية وتحديد سمات وسلوكات الأشخاص الذين افترضت بعبقريتهم

وفي الفترة التي ازداد الاهتمام بالأساليب الإحصائية لتحليل وتفسير نتائج الدراسات والبحوث في مجال علم النفس عامة ومجال الاختبارات التي ركزت على قياس الذكاء وتحديد مكوناته بشكل خاص ، برزت آراء تؤسس لنظرية الذكاءات المتعددة فقد طرح بيرت (Burt,1833) مفهوما جديدا للذكاء على انه موروث وأن الذكاء عبارة عن قدرة عقلية عامة وهذه القدرة هي التي تمكن الفرد من الكلام أو الحديث أو التصرف بطريقة واعية، وقد أشار إلى أن البيئة المحيطة بالفرد قد تؤثر في ذكائه لكن العامل الجيني أي الوراثي له الدور الأكبر في تشكيل الذكاء ، وكان يقصد بذلك الجينات الأبوية ، إضافة إلى أن المستوى الاجتماعي في مستوى ذكاء الفرد لكن يؤخذ على بيرت فيما بعد انه استخدم بيانات مزورة في سبيل إثبات وجهة نظره، وقد طرح فكرة تأثر الذكاء بعرق الشخص وجنسه، وبين أن الإناث يتمتعون بقدرات لغوية أعلى منها للذكور ، ومن أهم انجازاته تصميم اختبارات القبول في المدرسة عددها (١١) اختبارا للأطفال من سن الحادية عشرة وسمي باختبار ١١ + (Eleven plus) واستخدم في جميع انحاء بريطانيا .

وفي هذه المرحلة حاول ثيرستون (Therston,L,1887) تفسير الذكاء من خلال نظرية العوامل الأولية والتي تعتبر تأسيسا لنظرية الذكاء المتعدد إذ عمل على استخدام أسلوب التحليل العاملي في استخلاص مكونات الذكاء وتوصل إلى أن هناك سبعة عوامل تكون الذكاء منها القدرة اللغوية، الطلاقة، والقدرة المكانية والقدرة العددية، وسرعة الاستجابة، والقدرة المنطقية وقال بان الإجراءات الرياضية التي استخدمت في التحليل هي التي أنتجت عاملا واحدا. ولذلك عرف الذكاء على انه سمة عقلية تتكون من مجموعة من القدرات الأولية (primary abilities) التي تتناسق مما لتساهم في تميز شخص عن آخر وقال بان الذكاء قدرة على التفكير المجرد. ومما أكد وجهة نظره انه قام بتحليل درجات أشخاص متماثلين من حيث نسب الذكاء ومختلفين

من حيث قدراتهم وسلوكياتهم مما يعني أن العوامل التي تشكل الذكاء متعددة وليست مقتصرة على عامل واحد.

ومن المحاولات الجادة لدراسة الذكاء وتفسيره ومن ثم تصميم وبناء الاختبارات لقياسه كانت الطبيبة النفسية الأمريكية بيج كاتل (Psyche,Cattell,1893) وهي ابنة عالم النفس الأمريكي جيمس كاتل (J,Cattell,1860) حيث كانت تعمل في مختبر الصحة النفسية وكان من انجازاتها انها صممت اختبار ذكاء للأطفال القاصرين (Infant Intelligence scale) عام ١٩٤٠ ولقي إقبالا كبيرا لسهولة استخدامه، وانخفاض زمن تطبيقه وتركيزه على الجوانب المعرفية وعدم صلاحية اختبار بينيه الشائع آنذاك للأطفال صغار السن ، وهذا ما زاد من أهميته لدرجة انه أصبح يباع في عيادات الصحة النفسية، حيث استخدم في التنبؤ بذكاء الأطفال وتحديد خصائصهم النمائية رغم انخفاض معامل الارتباط بين النتائج عليه ونسبة الذكاء ، كما استخدم في تحديد خصائص النمو للأطفال ولذلك ساهم في تحديد بعض حالات التدخل المبكر مع الأطفال الضعيفين من ناحية النمو .

في هذه الأثناء بدت محاولات تصميم الاختبارات التي تقيس الذكاء خاصة مع تعدد تعريفات العلماء له حيث كانت أكثر المحاولات لبناء هذه الاختبارات من قبل وشسلر (Wechsler,1896) الذي عرف الذكاء على انه قدرة عامة تمكن الفرد من التفكير العقلاني والتكيف مع البيئة التي يعيش فيها.ويعتبر من أكثر مطوري اختبارات الذكاء فطور اختبار الذكاء للأطفال في العام (١٩٤٩) واختبار الذكاء للبالغين في العام(١٩٥٥) إضافة إلى تطوير اختباراته من قبل العديد من الباحثين والمهتمين فيما بعد، كما عزز مفهوم نسبة الذكاء الانحرافية واستخدم العمر العقلي في حسابها إلى جانب إجراء المقارنات المعيارية بين المفحوصين الصغار والكبار.

في العام ١٩٥٥ قدم عالم النفس الأمريكي جيلفورد (Guilford,1897) والذي كان رافضا لمفهوم الذكاء أحادي البعد وقال بأن الذكاء عبارة عن مجموعة من القدرات العقلية التي تمكن الشخص من معالجة المعلومات المختلفة أي انه يتكون من مجموعة من العوامل والتي قد تكون مستقلة عن بعضها واهتم بالفروق الفردية من

خلال اهتمامه بعلم النفس التجريبي، وفسر الذكاء من خلال زعمه أن الذكاء يتكون من مجموعة من القدرات العقلية التي تنتمي لثلاثة أبعاد هي بعد المحتوى وبعد العمليات وبعد النواتج وقدم نموذجه المعروف بمكعب جيلفورد للقدرات العقلية حيث كان قد تمكن من خلال عمله في سلاح الجو الأمريكي من قياس (٢٥) قدرة عقلية .وحسب هذا النموذج فان كل وجه من أوجه الكعب يمثل بعدا من أبعاد الذكاء الأول بعد المحتوى ويضم (٥) عناصر في حين تكون البعد الثاني العمليات من (٦) عناصر ، أما البعد الثالث فهو النواتج والذي تكون من (٦) عناصر وبهذا فان مكونات الذكاء تكون نتيجة للتفاعل بين عناصر الأبعاد الثلاث وبهذا يكون عدد القدرات العقلية حتى ذلك العصر هو(٥*٦*٥) أي(١٥٠) قدرة عقلية وهي عبارة عن تفاعل العناصر الفرعية في كل بعد من الأبعاد التي اقترحها جيلفورد

ومع ازدياد العلماء القائلين بتعدد القدرات المكونة للذكاء حاولت ثيلما ثيرستون (Thelma,Thurstone,1897) زوجة وليام ثيرستون الاستمرار بما جاء به زوجها وليام ثيرستون حيث كانت قد عكفت مع زوجها منذ العام ١٩٣٧ على بناء الاختبارات النفسية على مستوى التعليم العالي الأمريكي ، حيث كان يتم معايرة هذه الاختبارات سنويا وذلك حتى العام ١٩٤٨ كما عملت على تطوير إستراتيجية أخرى في التحليل العاملي كانت سببا في معارضة سبيرمان في تفسيره للذكاء بمفهوم العامل العام ، وقد انصب اهتمامها فيما بعد في العام ١٩٤٧ على إصدار أهم مجموعتين من الدراسات في هذا المجال الأولى تحت عنوان 'التعلم من اجل التفكير' في حين اهتمت المجموعة الثانية في 'القراءة للطلبة المتفوقين' كما اهتم عالم النفس التربوي فيليب فيرنون (Vernon,1905) بقياس الذكاء حيث عرفه على انه مركب من مجموعة من المكونات التي تساعد الفرد على التعلم والتكيف مع البيئة التي يعيش فيها ، كما أكد على أن لكل من العامل الجيني والبيئي اثر في تشكيل الذكاء وقال بان السلوك الإنساني هو نتيجة لتفاعل العوامل الوراثية مع العوامل البيئية .

وفي تلك المرحلة أكدت انستازي (Anastasi,1908) على تعدد العناصر أو العوامل المكونة للذكاء وقالت بأن الذكاء عبارة عن مركب من مجموعة من القدرات

الوظيفية التي تساعد الفرد على الحياة، وعلى الرغم من أنها تلقت تعليمها الابتدائي في البيت من قبل جدتها نتيجة لأسباب دينية إضافة إلى هروبها من المدرسة في التعليم الثانوي إلا أنها درست الرياضيات في بداية الأمر حيث اطلعت على ما قدمه سبيرمان في الارتباط والتحليل العاملي حيث أثارت أفكاره اهتمامها مما دفع بها إلى تحويل تخصصها في الجامعة من الرياضيات إلى علم النفس ، حيث اهتمت بدراسة الفروق الفردية وحاولت التركيز على العوامل المؤثرة فيها ، كما ركزت على اثر البيئة على درجات المفحوصين على اختبارات الذكاء ، لذلك اهتمت بالفروق الثقافية ومساهمتها في تحيز الاختبار لفئات دون أخرى وقالت بان الأطفال الذين يعيشون في بيئات فقيرة تتأثر درجاتهم على اختبارات الذكاء لذلك اعتبرت أن كثيرا من الاختبارات غير نقية بسبب معاناتها من عيوب التحيز الثقافي ، كما اعتبرت أن طبيعة القدرات التي يمتلكها المفحوصين مثل القدرة اللغوية والقدرة العددية والقدرات الخاصة تؤثر في درجاتهم وهذا ما يساهم في تصنيف هؤلاء المفحوصين في عدة فئات .

وتعتبر انستازي ممن قدموا مفهوما متطورا للاختبار حيث أكدت على أن الاختبارات التي تستخدم في ذلك الوقت تم بناؤها وفقا لتعريف أصحابها لسمة الذكاء إضافة إلى عدم مراعاتها للبيئات الاجتماعية حيث أنها تقيس قدرات الفرد في تلك البيئات بمعزل عن البيئات الأخرى، لان المفاهيم والخبرات السائدة في مجتمع ما تختلف عنها في المجتمعات الأخرى لذلك طالبت بان يؤخذ السياق الاجتماعي في الاعتبار عند تصميم وبناء الاختبار، إضافة إلى أن المهام التي قد تطلب من شخص ما قد تعتمد على قدرات لا تقيسها الاختبارات الحالية حيث تتغير متطلبات الحياة وهذا لا ينفي إمكانية الاستفادة من الاختبار بالتنبؤ بالأداء المستقبلي شريطة أن يبنى الاختبار بطريقة سليمة وقد بينت أن الاختبار يمكن أن يقيس قدرات ومهارات الفرد الخاصة بمهمات معينة وتحديد الأداء التعليمي والمهني المستقبلي وكذلك تطوير استراتيجيات الأداء وحل المشكلات، وانه إذا ما تم استخدام نتائج الاختبار بشكل صحيح سيتمكن المعلم من تحديد مستوى الأداء الأكاديمي الحالي للطالب، ولذلك نبهت انستازي إلى عدم منطقية التساؤل عن مقدار العلامة التي يحصل عليها المفحوص ولا بد من الاهتمام

بتخطيط التعلم والأنشطة الخاصة بناء على المستوى الذي يحدده الاختبار .وبناء على ذلك أطلق على انستازي اسم رائدة الاختبار بمفهومه الحديث (test guru) .

ومن الذين ساهموا في دراسة وتحديد القدرات المكونة للذكاء ثورنـدايك (Thorndike,1908) وهو ابن ادوارد ثورندايك الأب حيث واجه مشكلة تعدد مصادر الأخطاء وتحديد معايير اختيار منتسبي سلاح الجو الأمريكي ، الأمر الذي دفعه إلى البحث في آلية تحديد مواطن الضعف في الاختبارات المستخدمة آنذاك مما حدا به إلى استخدام التحليل العاملي في تحديد العديد من مصادر الأخطاء لاختيار الطيارين في سلاح الجو الأمريكي ووضع محك للاختيار ، كما اهتم بدراسة وتحليل مفهوم الثبات للاختبار وفي العام (١٩٥٤) قام بنشر مجموعة من اختبارات القدرات العقلية للمدارس حيث كان يهدف إلى استخدامها لتكوين ملف متكامل عن قدرات الطلبة في المدارس.

أما التطورات التي حصلت في المرحلة اللاحقة للقياس النفسي ولا سيما قياس القدرات العقلية للفرد فقد برزت في بدايات القرن العشرين من خلال العديد من مساهمات العلماء مثل عالمة النفس الأمريكية ريموند كاتل (Cattell,R.1905) وهي ابنة عالم النفس جيمس كاتل حيث جاءت إسهاماتها من خلال الطرق والأساليب الإحصائية التي اقترحتها وكذلك نظرية الذكاء المبلور ونظرية التأثير وهي إحدى النظريات التي تدرس الشخصية، فقد قالت بأن الذكاء عبارة عن مجموعة من القدرات المتنوعة التي تلتئم لتشكل عوامل متمايزة عامة وهذه القدرات قد تنتقل عبر الأجيال مثل سرعة التصرف والتذكر السريع والبعض الآخر يمكن اكتسابه بالخبرة مثل معرفة المعلومات والمعارف وقد يزيد عددها عن ال(١٠٠) قدرة منتشرة بين غالبية البشر ، وقد أشارت من خلال دراستها لنتائج تطبيق بعض اختبارات الذكاء أنَّ الأفراد الذين تكون نسبة الذكاء لديهم عالية يميلون إلى أن يكون عدد أطفالهم قليلا والأفراد الذين تكون لديهم نسبة الذكاء متدنية يميلون إلى أن يكون عدد الأطفال لديهم قليلا، ولذلك قال بأن من الحكمة أن يتم تشجيع الأفراد ذوي نسبة الذكاء العالية إلى ضرورة زيادة عدد أولادهم وتشجيع الأفراد من ذوي نسبة الذكاء المتدنية

على تقليل عدد أولادهم ، وقد حصل كاتل نتيجة لإسهاماته على جائزة المؤسسة الأمريكية لعلم النفس وكان الشخص رقم (١٣) الذي يحصل عليها منذ العام ١٩٥٦. بعد ذلك حاول جون كارول (Carroll,1916) تقديم صورة أوضح عن الذكاء وجاء بنظرية الذكاء المتعدد (الطبقي) Three-stratum حيث يتشكل ذكاء الفرد من مجموعة من العوامل العامة بين كل البشر ومجموعة من العوامل التي تختص بفئة أو مجموعة معينة من الأشخاص في حين أن هناك مجموعة من العوامل التي يتميز فيها الفرد ذاته ، وتعتبر هذه النظرية امتداد لنظرية العوامل المتعددة لكن الجديد فيها انه صنفها حسب عدد الأفراد الذين يمتلكونها. فالعامل العام يمتلكه الجميع والعامل الفئوي يمتلكه البعض دون الأخر والفريد يمتلكه فرد بعينه.

وفي هذه المرحلة ظهر عالم النفس الأمريكي ليون كامين (kamin,1924) واهتم في بداية الأمر في البحث في دراسة سلوك الحيوانات واعتبر من الباحثين المهتمين في هذا المجال ، وكان بذلك يريد الرد على القائلين بتفسير الذكاء بالوراثة وأشار إلى خطورة ذلك كما عارض آراء بيرت (Bert) في ذلك حيث كان يعتقد بتفسير الذكاء وراثيا ، وقد جاء اهتمام كامن بدراسة الذكاء متأخرا بعد عام (١٩٧٢) وذلك بعد حادثة مع طلابه في جامعة برنكتون (Princeton) تتلخص في دعوته لريتشارد هيرنستن (Herrnstein) - وهو احد المساهمين بوضع فكرة المنحنى الجرسي أحد نماذج منحنى التوزيع الطبيعي والذي كان قد نشر مقالا يدعم فيه تفسير الذكاء بالعامل الجيني (الوراثة) - وذلك للحديث إلى طلابه حول مضمون مقالته التي لم تروق للطلبة مما دفعهم إلى محاولة جر هيرنستن للحديث عن نسبة الذكاء الانحرافية (IQ) وقد تنبه إلى ذلك وحاول تجنب الإجابة مما اضطره إلى إنهاء محاضرته ، وهنا وجه الطلبة سؤالا إلى كامين فيما إذا اطلع على مقالة هيرنستن أو فيما إذا كان له موقف من مضمونها فما كان منه إلا الحديث عن موضوع الذكاء الذي كان موضوع جدل آنذاك، وسرعان ما أدرك أن آراء بيرت التي رفضها كانت تشكل خلفية مقال هيرنستن ، الأمر الذي جعله يشكك في نتائج دراسات بيرت وأصر على مراجعة تلك النتائج حيث كان كامن خبيرا إحصائيا آنذاك وذلك ليقارن بين ما جاء به بيرت وما

كتبه هيرنستن ، وحاول دراسة النظريات التي فسرت الذكاء بالعامل الوراثي حيث دعم التفسير البيئي للذكاء ولا زال كامين من العلماء المعاصرين

أما عالم النفس الأمريكي هوورن جاردنر (Gardner,1943) والذي عمل في طب الأعصاب في بداية حياته المهنية فقد قدم نظرية الذكاءات المتعددة في العصر الحديث وتحديدا في العام (١٩٨٣) حيث اعتبر أن الذكاء له أوجه متعددة وان كل نوع من الذكاء يتحدد بمضمونه أو محتواه ولذلك من أنواع الذكاء لديه الذكاء الاجتماعي، الذكاء الجسمي حتى انه اضاف ما سماه بالذكاء الانفعالي ، وعرف الذكاء على انه القدرة على حل المشكلات أو الإنتاج أو الابتكار ولذلك يعتبر من الرواد في تفسير الذكاء بنظريته المشهورة بالذكاءات المتعددة، وقد اعتمد في منطق نظريته على أن كثيرا من الأفراد يكون لديهم قدرات دون أخرى، فعلى الرغم من ضعف المعاقين في القدرة اللغوية لكن القدرة الجسمية أو الحركية لديهم قد تكون عالية وكذلك في التواصل إضافة إلى أن مصدر سمة الذكاء لا يقتصر على الجانب العقلي لوحده.وقد اقترح جاردنر سبعة أنواع من الذكاء (المكاني ، اللغوي، الموسيقي، الجسمي ، الموسيقي، الرياضي التواصل مع الآخرين والذكاء الشخصي)

وقد قدم عالم النفس المعرفي روبرت ستيمبرج (Robert J. Sternberg 1949) بعض التفسيرات التي اتفقت مع ما جاء به جاردنر وحيث ان ستيمبرج كان يعاني من تأثيرات اختبارات الذكاء المستخدمة في الدراسة المدرسية وإعطاء معلميه لتوقعات منخفضة عن أدائه المستقبلي لذلك أصر على دراسة الذكاء وإيجاد اختبارات بديلة لتلك التي كان يعاني منها ، لذلك كانت أول محاولاته هي بناء اختبار ذكاء موازي لاختبار بينيه في العلوم للصف السابع الابتدائي ، وأرسله إلى مدرسته التي كانت تبدي له التوقعات المتدنية، أشار إلى أن المؤشرات أو السلوكات التي تدل على الذكاء تبرز؟ أو تنتج من خلال عملية التوازن بين القدرات الخاصة بالتحليل والابتكار والقدرات العملية ، ولذلك عرف الذكاء على قدرة الفرد على النجاح في حياته الاجتماعية. والثقافية إضافة إلى قدرته على تحديد نقاط قوته وضعفه والقدرة على تنمية وتحسين

نقاط الضعف لديه ومن هنا سميت نظريته بنظرية النجاح، ولذلك اعتبر قدرة الفرد على اختيار المهنة او مجال العمل في المستقبل مؤشرا على الذكاء.

وفي هذه الفترة أيضا ظهرت اهتمامات عالم النفس الأمريكي كوفمان (kaufman,1944) والذي نشر كتابا مع مشلر بعنوان اختبارات الذكاء "intelligent testing" والذي انتشر وعرف بين الناس، كما عمل على تطوير اختبارات وسشلر لذكاء الأطفال (K-ABC) وكذلك اختبار التحصيل التربوي الذي عرف باسم(K-TEA) والموجز في اختبارات الذكاء (K-BIT) وكذلك اختبارات الذكاء للمراهقين والكبار واختبار كوفمان للتحصيل الأكاديمي Kaufman Test of Educational Achievement (K-TEA), وكذلك اختبار كوفمان للذكاء (Kaufman Brief Intelligence Test (K-BIT)، كما تم إخراج الطبعة الثانية من هذه الاختبارات فيما بعد (KABC-II, KTEA-II, KBIT-2) .كما قدم سلسلة من الأدلة حول تقييم الكتب تحدث من خلالها عن العديد من أدوات القياس هذا الاختبارات وكتب العديد من المقالات حول الاختبارات وأشار إلى الاختبارات المقالية، ونتيجة لذلك أصبح كوفمان خبيرا ومستشارا في إعداد الاختبارات بمختلف أنواعها .

وفي السبعينات من القرن الماضي حاولت عالمة النفس الامريكية كارول دويك (Dweck,1946) الوقوف على مفهوم الأفراد من حيث طبيعة ذكاءهم الشخصي وقالت بان هناك نمطين من الأفراد فهناك الأفراد الذين يؤمنون بان ذكائهم كامن وهو ثابت وهو من السمات غير القابلة للتغيير أي أن الذكاء سمة ثابتة ، أما النمط الأخر فيعتقدون بان الذكاء سمة متطورة قابلة للزيادة أو التطور ، وبهذا فلم تقدم دويك تعريفا محددا للذكاء إنما حاولت أن تدرس وتفسر مفهوم الأفراد له وتدرس سلوكات الأفراد كل حسب النمط الذي ينتمي إليه. فالافراد الذين يعتقدون بان الذكاء سمة ثابتة وداخلية لوحظ ان قدرتهم على تحدي المهمات التي تواجههم قليلة كما أنهم يشعرون بالتهديد والخطر نتيجة فشلهم الأكاديمي، أما بالنسبة للأفراد

الذين يعتقدون بان الذكاء سمة قابلة للتطور فان تحفيزهم ومدحهم يقلل او يحد من تطوير سمة الذكاء لديهم وقد قدمت مبررا لهذا الادعاء.

حاولت دويك دراسة الحوافز الإنسانية خاصة مع الأفراد بطيئي الاستجابة (Deiner & Dweck, 1978, 1980; Dweck, 1975; Dweck & Reppucci, 1973). ولاحظت أن بعض الطلبة الذين أجرت عليهم دراساتها يبقون مستسلمين للشعور بالفشل الدائم بينما البعض الآخر تجاوز هذا الشعور واخذ الأمر على محمل الجدية واستمر بالمحاولة من أجل النجاح، وفي عام (١٩٨٠) حاولت البحث في المفهوم النظري الذي يكمن خلف السلوكات التي لاحظتها ، لتكتشف أن المعتقد الذاتي حول طبيعة الذكاء له تأثير كبير في تحديد المنهجية المتبعة في حل أو انجاز المهمات أو السلوكات فالطلاب الذين يعتقدون بأن ذكائهم لن يتغير يميلون إلى الخجل من التحديات خاصة في المجال الأكاديمي بينما الطلبة الذين يعتقدون بقابلية تطور الذكاء فان ذكاءهم يزداد بالاجتهاد والمحاولة والمثابرة . وقد وجدت بالتجربة أن الطلبة الذين يعتقدون بثبات ذكاءهم يتفقون مع القول بان ذكاءهم لن يتطور كثيرا ولذلك فهم يبدون قيمة عالية للنجاح ، إضافة إلى شعورهم بالقلق إذا ما تطلب الأمر منهم العمل بجد لاعتقادهم بانخفاض مستوى قدراتهم. لذلك فان خياراتهم الأكاديمية تتمحور حول الأشياء التي يمكن إن يؤدوها بشكل جيد ، ومن هنا فان خياراتهم تكون ذات مستوى متدني ليبدو أداؤهم فيها عاليا ، وفي المقابل فان الطلبة الذين يعتقدون بإمكانية تطور الذكاء فإنهم لا يشعرون بالفشل أو القلق لاعتقادهم بان قليلا من الجد والمثابرة يزيد من قدرتهم على التحدي والنجاح. وقد أشار ت دويك إلى أن الثناء والتحميز الذي يبديه الكثير من المعلمين وأولياء الأمور على الطلبة قد يدفعهم خالي تبني الاعتقاد بثبات الذكاء وعدم إمكانية تطويره. كما أشارت إلى أهمية الكلمات العبارات المستخدمة في الثناء والتحفيز فالطالب الذي يقول له والده أنت ستنجح مثل زميلك الذكي قد يخشى أن يفهم أن فشله يعني انه غبي ولذلك ركزت على أهمية اختيار عبارات الثناء على الطلبة، حيث أن ذلك له علاقة بتطور مفهوم الطالب عن ذاته.

وأخيرا فقد اعتبر عالم النفس الأمريكي ايث سيمونتون (Simonton,1948) أن الذكاء عبارة عن مجموعة من القدرات العقلية إلي تساعد الفرد على التكيف الاجتماعي الناجح وطرح أمثلة عليها مثل القدرة على استرجاع المعلومات والقدرة على حل المشكلات ، وقد اتفق مع جاردنر (Gardner,1943) في مفهوم الذكاء الاجتماعي كنوع من أنواع الذكاءات المتعددة وليس غريبا عليه هذا التفسير فهو متخصص في علم النفس الاجتماعي.

وقد أهتم سيمونتون بدراسة العبقرية والسير الذاتية للعباقرة ونشر العديد من المقالات ولكتب في هذا المجال، حيث أراد أن يستقصي جملة من المبادئ العامة حول السلوك الإنساني من خلال سلوكات العباقرة والمبدعين أو ما سماهم بالشخصيات المهمة. ومن هذه الاستنتاجات:

- انه وبالرغم من أهمية نسبة الذكاء (IQ) في التفوق الأكاديمي إلا أن التحليلات الإحصائية اثبتت أن مساهمة العنصر الثقافي في ذلك لا تتجاوز (٤ – ٥ ٪) وان الجوانب التطورية والعوامل الشخصية والتحفيز له الدور الأكبر في الانجاز ، ولذلك فهو يرى أن الشعوب الذكية ليس بالضرورة أن تحقق انجازات عظيمة (Simonton, 1999a & Simonton 2003)
- ومن استنتاجاته أيضا أن العلاقة بين الذكاء والقناعة مهمة في التأثير بالآخرين حيث أن التأثير بالآخرين مهم باعتباره مؤشر على التواصل والذي بدوره يعتبر سلوكا دالا على الذكاء بمفهومه الاجتماعي ، فالذكاء يساهم في تقليل الجهد والوقت للوصول إلى قناعات ورغبات الأشخاص، (Simonton, 1999a) كما استنتج أن مستوى الذكاء له دور مهم في الإبداع في الحياة مثل حل المشكلات ، لكن الأشخاص المبدعين ليس بالضرورة أن يقدموا إبداعات واضحة وقد ينطوي ذلك على المجتمع بشكل عام، (Simonton, 1999a) .

ومن خلال هذا العرض يتبين أن تطور علم القياس برز من خلال الاهتمام بدراسة الشخصية الإنسانية وقد تركز هذا الاهتمام على دراسة قدراته التي تمثلت في بداية الأمر بقدراته العقلية ولذلك اقترنت الكثير من التطورات والمفاهيم لهذا العلم

بالقدرات العقلية ولذلك فان اهتمام العلماء بفهم وتفسير سمة الذكاء أو القدرات العقلية ساهم إلى حد بعيد في تطور مفاهيم القياس وإجراءاته وكذلك أساليبه الإحصائية التي كانت متطلبا لتحليل البيانات التي كان يتم الحصول عليها نتيجة للدراسات والاختبارات التي كانت تطبق كممارسات بحثية لفهم الذكاء الإنساني ، وهذا بحد ذاته ساهم في تطور أدوات القياس ولا سيما الاختبارات العقلية ، ومن ثم توالت الاهتمامات بالأدوات التي استخدمت بدراسة السمات الإنسانية الأخرى كالدافعية والاتجاهات والتي عرفت باسم المقاييس كمقاييس الاتجاهات والميول المهنية والمسوحات الاجتماعية والمقاييس الاجتماعية ومقاييس الأحكام التفاضلية لليكرت والتي نتج عنها قوائم الشطب وسلالم التقدير ومقاييس التباين اللفظي والملاحظة.. .الخ.

ومع تزايد الاهتمام ببناء الاختبارات وضرورة الدقة في المعلومات التي يتم الحصول عليها من جراء تطبيق الاختبارات ، وبروز خصائص الاختبارات والمقاييس مثل الثبات الصدق ومؤشرات الصعوبة والتمييز وتحليل الدرجات عليها بدت معالم نظرية القياس تبرز إلى حيز الوجود من خلال الإحصائيات الخاصة بالاختبار باعتباره أداة قياس الأمر الذي تطلب توظيف المفاهيم الإحصائية في توفير المعلومات حول دقة القياس ، فبرز مفهوم الخطأ المعياري للقياس والدرجة الحقيقية والملاحظة ووحدة القياس لمختلف المتغيرات الأمر الذي أدى إلى تصنيف المتغيرات حسب طبيعتها إلى مستويات القياس المعروفة وبرزت إلى حيز الوجود ملامح نظرية القياس التي باتت تعرف الآن بنظرية القياس التقليدية وبعد ذلك برز إلى حيز الوجود اعتبارات أخرى للاختبارات والدرجات والقدرة حيث برز في بداية الثمانينات فيما عرف بنظرية القياس الحديثة.

وخلاصة القول ولان الهدف من هذا الكتاب طرح القياس من المنظور الكلاسيكي والحديث كان لا بد من الحديث عن التطور التاريخي لإجراءاته وعملياته والتي اقترنت بتطور دراسة الخصائص الإنسانية ولا سيما القدرات العقلية ، حيث تضفي هذه المقدمة التاريخية المزيد من التهيئة لفهم افتراضات النظريتين وتطورها.

الفصل الثاني

مفاهيم القياس

مقدمة

يعد القياس والتقويم التربوي أحد المكونات الرئيسة للمنظومة التربوية، وأكثرها تأثيراً في تقدمها وازدهارها، إذ يسهم تطويرها إسهاماً إيجابياً في إصلاح وتطوير المكونات الأخرى لذلك فقد حظي مجال تطوير نظم وأساليب القياس والتقويم باهتمام كبير من قبل المهتمين بتطوير النظم التعليمية استناداً إلى ما أدت إليه البحوث والدراسات التربوية المعاصرة من حيث الاهتمام بعمليات وأساليب القياس المتطورة في توجيه مسار العمل التعليمي والنهوض به، وتحديد مدى تحقيق النظام التعليمي لأهدافه المرجوة .وتعد الاختبارات والمقاييس النفسية أحد الوسائل الهامة والضرورية لعمليات التقويم التي تعتمد عليها المؤسسات التعليمية والتربوية للتعرف على مدى التقدم في تحصيل الطلاب ومدى تحقيقهم للأهداف التعليمية، لذلك يمكن اعتبارها أساسا هاما في صنع القرارات المصيرية لكل من تطبق عليه هذه الاختبارات .وبالتالي فإن دقة وسلامة تلك القرارات تتوقف على مدى جودة وكفاءة أدوات القياس المختلفة ومنها الاختبارات المستخدمة في عملية التقويم والأساليب المستخدمة في تحليل وتفسير نتائجها (علام، ١٩٨٧)

إذ أن تقويم الطلبة يعتمد على تحليل البيانات التي يتم الحصول عليها عن طريق الاختبارات وأدوات القياس المختلفة، ويعتمد على كافة البيانات الأخرى التي تساعد على معرفة التغيرات التي تطرأ على نمو الطالب في جميع جوانب شخصيته .ومن هنا تظهر أهمية سلامة البيانات التي يتم جمعها والتقويم في ضوئها، وبالتالي التأكد من كفاية الأدوات التي تستخدم في جمع هذه البيانات .(القرشي، ١٩٨٦)

وبالرغم من الإيمان المسبق بوجود خطأ في تقدير الدرجة التي يحصل عليها المفحوص في اختبار ما بالزيادة أو النقصان، فإنه يجب ألا يصل الخطأ إلى درجة تصبح عندها هذه الدرجة مضللة لكل من المفحوص طالبا أو موظفا أو للمدرس أو ولي الأمر وكل من تهمه الدرجات . فقد يشعر أحد المفحوصين بأن الدرجة التي حصل عليها تخفض من مستوى تحصيله الحقيقي في حين يشعر آخر بأنها تبالغ في زيادة مستوى تحصيله الحقيقي، إلا أنه بحكم رغبته في درجة عالية قد يبدي اتجاهها إيجابياً نحو المقرر ومدرسه وحتى نحو نوعية الفقرات التي يجيب عليها، ولكن سرعان ما يكتشف أن درجته مجرد رقم مضلل عندما يخضع لمحك معين، وعندها قد يبدأ اتجاهه الإيجابي بالتحول (عودة، وحوامدة، ١٩٩٦).

علم القياس :

يتعلق علم القياس بتوفير المعلومات وهذا يتطلب الاهتمام بأدوات القياس وإجراءات ضمان دقتها وإجراءات بناءها أو تصميمها أو حتى تطويرها أي تكييفها من بيئة إلى أخرى أو من فئة لفئة أخرى ويمكن تعريف القياس في أنه فرع من فروع العلوم التطبيقية وأقرب ما تكون إلى الاحصاء التطبيقي يحاول وصف وتنظيم وتقييم نوعية القياسات أو جودتها و العمل على أن تكون القياسات أكثر دقة وأكثر فائدة وذات معنى ، إضافة إلى اقتراح أساليب لتطوير أدوات القياس .وهذا يعني أن لهذا العلم أو الحقل إجراءاته وعملياته ومفاهيمه والتي سوف نبدأ بالعرض لها.

وتميز أدبيات القياس النفسي بين مدخلين رئيسين في بناء الاختبارات والمقاييس وتقييم وتحليل البيانات المستمدة منها، وهذين المدخلين هما المدخل الكلاسيكي أو التقليدي والمدخل المعاصر .والمدخل الأول يعتمد على النظرية الكلاسيكية Classical Test Theory وما تنطوي عليه من مفاهيم ومبادئ بعضها يتعلق بخصائص مفردات الاختبار كالصعوبة والتمييز والبعض الآخر يتعلق بخصائص الاختبار ككل مثل الصدق، الثبات، والمعايير .

أما المدخل الثاني فيعتمد على النظرية المعاصرة للاختبارات Modern Test Theory والتي تعرف باسم نظرية السمات الكامنة Latent Trait Theory (LTT)

أو نظرية استجابة على الفقرة (IRT) (علام، ١٩٩٠) ولكل من المدخلين طرقه الخاصة في تقدير قدرة الفرد الحقيقية من الدرجة وهذه الطرق تختلف في المدخلين الكلاسيكي أو التقليدي والمعاصر أو نظرية السمات الكامنة، وسيقدم هذا الكتاب علم القياس من حيث المفاهيم والعمليات والإجراءات من خلال منظور النظرية الكلاسيكية ومنظور النظرية الحديثة ، ولكن في البداية لا بد من التقديم للمفاهيم الأساسية في القياس.

مفاهيم أساسية في القياس

يتضمن القياس العديد من المفاهيم والتي تشكل بمجموعها إجراءات وأدوات القياس ، حيث أن الوعي بهذه المفاهيم من شأنه أن يسهل على القارئ الوعي بطبيعة القياس وتفسير عملياته ونتائجه وسوف نعرض لهذه المفاهيم كمقدمة لموضوعات هذا الكتاب.

الاختبار Test

يعتبر الاختبار من ابرز مفاهيم القياس ، فلو تأملنا عن عدد الاختبارات التي خضعنا لها خلال حياتنا الدراسية أو الوظيفية ، بلا ادنى شك سيكون الرقم كبيرا وإن ابسط تعريف للاختبار هو "إجراء منظم لقياس عينة من السلوكات من خلال عينة من المثيرات" أي انه أداة للحصول على معلومات حول سلوك الفرد ونقول عينة لأنه لا يمكن الحصول على جميع سلوكات الفرد وبالتالي لا يمكن تحديدها وبالتالي إخضاعها للاختبار وتصمم الاختبارات عادة لأغراض متعددة .

التقويم Evaluation

التقويم يعني دائما اتخاذ القرار اعتماداً على تجميع المعلومات وتفسيرها فإعطاء التلاميذ درجاتهم يعني النجاح أو الفشل واختبار الأطباء يعني إجازتهم في الطب وكذلك اختبار متقدمين لاختبار قيادة السيارة مثلا يعني نجاحهم أو رسوبهم وبالتالي منحهم الرخصة أو عدم منحها .

الاختيار : Selection

يتم الاختيار للأفراد إما للتوظيف أو الدراسة فكثيرا ما تعلن الشركات عن توافر عدة وظائف حيث تقوم باختبار المتقدمين لتختار من بينهم من يقوم بملء هذه الشواغر أو الوظائف، فالاختبار الذي تقوم به الشركة يكون الغرض منة أن تختار الحاصلين على الدرجات العليا مثلا أو أن تحدد مستوى معين وتقوم الشركة بتعين من حقق هذا المستوى وبلغة القياس نقول قد نعتمد طريقة معيارية أو محكية المرجع، ويندرج تحت هذا الغرض أيضا امتحانات القبول في الجامعات فالمتقدم قد يختار وقد لا يختار قد يقبل أو لا يقبل" .

التصنيف والتعين : Placement and Classification

عادة ما يتم تصنيف الأفراد كالطلبة أو الموظفين إلى فئات أو مستويات ولذلك لا بد من توفر معيار يتم التصنيف اعتمادا عليه وخير مثال على ذلك اختبار المستوى الذي يجرى في الجامعات لتصنيف مستويات المتقدمين مثلا في اللغة الانجليزية ويتم بموجبة تصنيف المتقدمين إلى مستويات تبعا لأسس محددة

الإرشاد : Counseling

الإرشاد هنا بمعنى التوجيه إلى مسار دراسي أو مجال عمل معين اعتمادا على معلومات حول من يراد لهم التوجيه كالطلبة أو المتخصصين ..الخ وأمثلة ذلك كثيرة منها اختبارات الإرشاد من اجل اختيار مهنة أو دراسة تخصص معين أو الإرشاد الخاص بعملية الزواج كما يتم في المجتمعات الغربية

القياس :

القياس يعني تعين (إعطاء) أعداد أو رموز للأفراد أو الأشياء بطريقة منظمة كوسيلة لتمثيل خصائص الأفراد وتعين الأعداد للأفراد يتم حسب قواعد موصوفة بدقة ، وإذا دققنا في تعريف القياس ستستوقفنا عبارة طريقة منظمة ، حيث توحي لنا بالتعليمات الواحدة والإجراءات الواجب إتباعها إذ لا يمكن مقارنة العلامات بشكل ذي معنى إذا اختلفت التعليمات أو فقرات الاختبار أو طرق التصحيح فإذا أعطينا

الأشخاص ذوي العيون الزرقاء العدد (١) وأعطينا الآخرين الرقم (٢) فان ذلك قياس لان الأعداد أعطيت للأفراد بشكل منتظم كما أن الفروق بين الأعداد تمثل فروقا في الخاصية وهي لون العيون كما أن اختلاف الأعداد الناتجة عن عملية القياس تعني دائما اختلاف في السمة المقاسية .

مستويات القياس :

يمكن أن يحدث القياس في اربع مستويات مختلفة

١ ـ الاسمي Nominal مثل تقسيم الصف إلى شعب معينة ' إعطاء الشوارع أرقام

٢ ـ الرتبي Ordinal إعطاء الطلاب رمزا من الحروف في امتحان ما .

٣ ـ الفتري Interval قياس درجة الحرارة والرطوبة والضغط الجوي .

٤ ـ النسبي : Ratio مقاييس الطول والوزن والحجم والمسافة

ويحدد كل مستوى من المستويات الأربعة كيفية ارتباط الأعداد المعطاة للأفراد بالسمات المعينة ، كما ويوفر مستوى القياس واحدة أو أكثر من أربع وظائف هي :

١ ـ التصنيف : أو التميز Classification and Discrimination

٢ ـ ترتيب الأشياء أو الأفراد حسب مقدار السمة Ordering In Magnitude

٣- الفئات المتساوية Equal Intervals.

٤- الصفر المطلق Absolute Zero.

ويبين الجدول رقم (١) مستويات القياس وخصائصها التي تتمتع بها بشكل هرمي من الأدنى إلى الأرقى من حيث الدقة وهو ما يعرف عادة بمفهوم هرمية المقاييس .

جدول :(١) مستويات القياس وخصائصها والتي توضح هرمية المقاييس .

	تصنيف	ترتيب	فئات متساوية	الصفر المطلق
تصنيف	نعم	نعم	نعم	نعم
ترتيب	لا	نعم	نعم	نعم
فئات متساوية	لا	لا	نعم	نعم
الصفر المطلق	لا	لا	لا	نعم
العمليات المسموح بها	أي تحويل بشرط إعطاء أعداد مختلفة	تحويل طردي خطي أو غير خطي	الاقتران الخطي فقط وعلى صورة ص = أس + ب بحيث أن أ ≠ صفر	الضرب والقسمة ولا يسمح بالجمع والطرح

● ترجم عن كتاب introduction totest thory (Groger and algena, 1986)

نلاحظ من خلال الجدول أن السمة تكون سمة هرمية القياس والهرمية تعني أن كل مقياس يتضمن خصائص المقاييس التي تسبقه فمقياس النسبة يتضمن خصائص المقياس الفئوي والرتبي والاسمي وللتأكد يمكن ترتيب كلمة نعم فنلاحظ انه ينتج لدينا هرم له بداية ونهاية .

أمثلة:

المقياس الاسمي:

ويعتبر هذا المقياس من ابسط مستويات القياس:مثل إعطاء الذكور في متغير الجنس الرقم (١) أو الحرف (أ) والإناث رقم (٢) أو الحرف (ب) مع ملاحظة أن هذا لا يعني أن (٢) أكبر من (١) أو ان (أ) افضل من (ب) أي ان الأعداد أو الرموز هنا لا تحمل المعنى الكمي مطلقا .

مثال (١) . إعطاء الأشخاص أعدادا حسب مؤهلاتهم العلمية على النحو التالي

١. بكالوريوس.
٢. بكالوريوس + دبلوم
٣. ماجستير.

مثال (٢) : إعطاء الأشخاص أعدادا حسب جنسيتهم على النحو التالي.

١. أردني.
٢. سوري
٣. عراقي

صحيح أن الفروق بين الأعداد هنا متساوية والفرق هو (١) لكن هذا الفرق لا يعكس فروقا غير متساوية وهي نفس الفروق بين الذكور والإناث وبين العراقي والسوري والأردني من حيث الجنسية بل تعكس فروقا غير متساوية وهي نفس الفروق بين الذكر والأنثى وهذا الفرق يختلف عن الفرق بين البكالوريوس والماجستير والدكتوراه أي لا يوجد منطق لترتيب معين ويمكن أن تترتب عدة ترتيبات ولا تتغير

المقياس الرتبي:

وهو المستوى الذي يمكننا من ترتيب أفراد المجموعة تنازليا أو تصاعديا حسب درجة امتلاكهم لسمة معينة كأن يتم ترتيب أفراد المجموعة حسب مستويات الذكاء لديهم أو قدرتهم على الانتباه ، وهكذا نرى أن القياس بهذا المستوى لا يكتفي بأن يبين أن الأفراد يختلفون بالنسبة لسمة معينة كما هي الحال في القياس الاسمي بل ويرتبهم أيضا حسب درجة امتلاكهم لهذه السمة ، لاحظ أن كم المعلومات يزداد بانتقالنا إلى مستوى قياس أرقى أو أدق والدقة أو الرقي تأتي من زيادة كم المعلومات ومثال ذلك ترتيب الطلبة حسب الصفوف أو المرحلة التعليمية التي ينتمون إليها.

١ الصف الأول.	١ التعليم الابتدائي.
٢ الصف الثاني	٢ التعليم الإعدادي
٣ الصف الثالث.	٣ التعليم الثانوي

المقياس الفئوي :

وهنا تزداد أيضا المعلومات التي نحصل عليها في هذا المستوى وتبعا لذلك تزداد الدقة و خير مثال على ذلك درجات الحرارة موزعة بوحدة الخمس درجات من الصفر إلى عشرين أو علامات طلاب صف تتوزع بين الصفر والمائة بوحدة العشر نقاط .

الفئة الأولى (صفر– ٥)

الفئة الثانية (٦– ١٠).

الفئة الثالثة (١١– ١٥).

الفئة الثانية (١٦– ٢٠).

لاحظ أن لدينا معلومات تشمل

- اسم الفئة (الأولى ، الثانية ، الثالثة) وهذا مستوى قياس اسمي.
- الفئة الأولى اقل من الثانية والثالثة والفئة الثانية أقل من الفئة الثالثة من حيث درجة الحرارة أو العكس أي أننا يمكن ترتيب هذه الفئات حسب درجة الحرارة من الأعلى للأقل أو العكس وهذا مستوى قياس رتبي.
- الفروق بين الفئات فمثلا الفرق بين الفئتين الأولى والثانية ٥ درجات حرارة وبين الثانية والثالثة كذلك . وهذا مستوى قياس فتوي.

المقياس النسبي

وهنا يكون القياس دقيقا إلى درجة كبيرة بسبب توفر أداة دقيقة وواضحة للقياس وكذلك أن السمة تكون محسوسة مثل قياس الطول بوحدة قياس الطول وهي المتر ومشتقاه ،وقياس درجة الحرارة ، بوحدة الكلفن أو الفهرنهيت. . وأهم خاصيتين لمقياس النسبة هما خاصية الصفر المطلق أي النقطة الدالة على انعدام السمة ومثال على ذلك إذا أردنا قياس طول مجموعة من اللاعبين فيمكن إعطاء كل مجموعة اسما أو ونرتبهم ونصنفهم في فئات ونجد الفرق الحقيقي بين كل فئة وبين كل لاعب في الفئة الواحدة وذلك على النحو التالي:

١ أو أ للمجموعة من ذوي الطول العالي (١٧٠ إلى ١٩٠)

٢ أو ب للمجموعة من ذوي الطول المتوسط(١٥٠ إلى ١٧٠).

٣ أو جـ للمجموعة من ذوي الطول المتدني (١٣٠ إلى ١٥٠).

لاحظ أن قياس سمة الطول يمكننا من إعطاء كل لاعب اسما ويمكن أن يأخذ اسم الفئة وهذا قياس اسمي.إعطاء كل لاعب رتبة حسب طوله فاللاعب الذي طوله(١٥٥ سم) يون اقل رتبة من لاعب طوله (١٦٥) أي ان طول اللاعب يحدد رتبته ، وهذا قياس رتبي.

ويمكن أن نصنف هؤلاء اللاعبين في فئات وكل فئة تختلف عن الفئة التالية لها أو السابقة لها مباشرة بمقدار (٢٠سم) وهي فروق متساوية .

ويمكن ايضا لتحديد الفروق بين اي لاعبين في أي فئة وفي الفئات الأخرى لوجود الصفر المطلق ووحدة القياس والسمة المحسوسة.ويمكن توضيح العمليات الحسابية المسموح بها في مستويات القياس :

الاسمي :

يسمح باستخدام جميع العمليات الحسابية مادامت تعطي اعدادا مختلفة

مثال : إذا أعطي الطلاب الذكور الرقم (١) والإناث رقم (٢) وإذا زدنا (٢) لكل رقم فأن الناتج سيكون . ١+٢=٣ وكذلك ٢+٢=٤

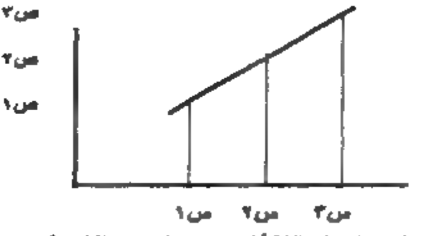
لاحظ اختلاف الأرقام ٣و٤ وكذلك إذا ضربنا كلا منهما بالرقم (٥)

١×٥=٥ ٢×٥=١٠ ومثل ذلك في الطرح والقسمة .

الرتبي : يسمح بأي تحويل طردي (أي يحافظ على الترتيب) خطي كان أم غير خطي ويخرج من ذلك القيم السالبة

الفئوي :

يسمح باستخدام العمليات الحسابية الأرجح ولكن بشرط ألا يكون على شكل التحويل الخطي. وهذا يعني أن فروقا متساوية بين الأرقام في محورا لسينات تقابل فروقا متساوية في محور الصادات كما في الشكل رقم (١).



شكل (١) التحويلات الخطية التي يسمح بها مستوى القياس الفتري

النسبي :

يسمح هنا بالضرب والقسمة ولا يسمح بالجمع والضرب وكما هو موضح بالعلاقة ص=أ.س فإذا كانت ص =٢س فأن التحويلات الممكنة تكون على الشكل التالي:

س	ص	
١	٢×١=٢	
٢	٢×٢=٤	
٣	٢×٣=٦	وهكذا في القسمة

المتغير والثابت Constant & Variable

توجد السمات والخصائص في الطبيعة بشكل متغير غير ثابت وهنا نسميها متغيرات أما الثابت فهو قيمة لعدد محدد غير متغير على عكس المتغير الذي يأخذ قيما عددية مختلفة وهنا قد يطرح السؤال التالي هل الثابت متغير ؟ الجواب طبعا لا لان الثابت له قيمة واحدة فقط وحتى يصبح متغيرا يجب أن يكون له قيمتان على الأقل . وهناك نوعان من المتغيرات : متغيرات متصلة ومتغيرات منفصلة والفرق بينهما أن

العدد في المتغيرات المنفصلة يأخذ قيمة محددة مثل عدد المدارس أو عدد المعلمين أو عدد الشعب للصف الواحد ، أما المتغيرات المتصلة فهي تأخذ اي قيمة في مدى معين ضمن حدود الدقة التي يسمح بها القياس مثل التحصيل والذكاء والطول والوزن اي انه يأخذ الوحدات وأجزائها بمعنى انه يحتمل الكسور والأعداد العشرية.

مفاهيم اساسية في الإحصاء

يستخدم علم القياس في إجراءاته العديد من المفاهيم الاحصائية بحيث يتم توظيفها للحصول على المعلومات كالتحليل الإحصائي للبيانات وكذلك تفسير نتائج التحليل من اجل فهم عمليات القياس وتفسير الظواهر والتنبؤ بها وصياغة التقارير حسب المستخدمين ولمختلف المفاهيم والعمليات الإحصائية المستخدمة من حيث الغرض والتعقيد ، ومن المهم هنا أن نميز بين نوعين من الاحصاء هما الإحصاء الوصفي والإحصاء الاستدلالي وذلك من العرض التالي

الإحصاء الوصفي :

يشير مفهوم الاحصاء الوصفي إلى المؤشرات الإحصائية البسيطة التي تصف المشاهدات وإحصائياتها الأولية ويشار إليها أحيانا بالموصفات (Destructives) ، حيث انه من المعروف أن أي مجموعة من المشاهدات حول أي ظاهرة تنتظم في توزيع معين ويقصد بالتوزيع الخصائص الإحصائية التي تتلخص بشكل التوزيع أي كيفية انتشار العلامات والأوصاف الإحصائية التي تلخص هذه المشاهدات، كالوسط الحسابي والانحراف المعياري واكبر مشاهدة وأقل مشاهدة .الخ وهناك مجموعة من المفاهيم التي تندرج ضمن الاحصاء الوصفي وتلتئم مع بعضها بطبيعة محددة والتي لا بد من الحديث عنها ضمانا لفهم عمليات وإجراءات علم القياس وسنتعرض لكل واحدة منها ، وهي على النحو التالي:

أولا: مقاييس النزعة المركزية Central Tendency.

سميت بهذا الاسم لأنها تعبر عن وسط التوزيع ، أي أنها جميعها قيم تأتي من وسط التوزيع بل تنزع إلى أن تكون في وسط التوزيع وسيتضح ذلك عن الحديث عنها وهي :

الوسط الحسابي Mean

وهو الأكثر شيوعا كمقياس للنزعة المركزية ويحسب عادة من خلال قسمة مجموع المشاهدات على عددها ويأتي في وسط المشاهدات بعد ترتيبها ترتيبا تصاعديا أو تنازليا وهو ممثل لجميع المشاهدات أي يعبر عنها مجتمعة وهو ثابت نسبيا على الرغم من كونه الأكثر تأثرا من المنوال والوسيط بالقيم المتطرفة ويعرف حسابيا بالمعادلة رقم(١)

$$\text{س} = \frac{\text{مجموع قيم س}}{\text{عدد القيم}} \quad \ldots\ldots\ldots\ldots(١)$$

المنوال Mode

وهو القيمة الأكثر شيوعا أو تكرارا وغالبا ما يكون في وسط المشاهدات.

الوسيط: Median

وهو المشاهدة التي تقسم المشاهدات إلى قسمين متساويين وهو اقرب إلى الوسط الحسابي من المنوال وسنأخذ المشاهدات التالية لتوضيح فكرة النزعة المركزية

مثال:

إذا كانت علامات عشرة طلاب تتوزع على النحو التالي احسب مقاييس النزعة المركزية لتوزيع العلامات.

٤ ، ٣، ٥، ٤، ٤، ٧، ٩، ٦ ، ٨، ١٠، لا بد من ترتيب العلامات تصاعديا أو تنازليا على النحو التالي:٣، ٤ ، ٤، ٤ ، ٥ ، ٦ ، ٧ ، ٨ ، ٩ ، ١٠

الوسط الحسابي

هو مجموع الدرجات مقسوما على عددها مجموعها = ٦٠ وعددها ١٠ إذا الوسط الحسابي = ٦٠/ ١٠= ٦ لاحظ أن القيمة (٦) تأتي في وسط التوزيع

الوسيط :

بعد ترتيب القيم نعد القيم الأقل الوسيط والأكثر منه ويجب أن تكون متساوية والقيم التي تحقق هذا الشرط هي القيم (٥ ، ٦) وهنا نجمعهما ونقسم على ٢ أي ٥ +٦ = ١١ / ٢=٥,٥ لاحظ أن القيمة من وسط التوزيع وقريبة من الوسط (٦)

المنوال لاحظ أكثر قيمة مكررة وهي القيمة (٤) وهي أيضا من وسط التوزيع

لاحظ أن جميع قيم النزعة التي أوجدناها تنزع إلى أن تكون من وسط التوزيع ولذلك جاء اسمها كذلك.

ثانيا: مقاييس التشتت Measures of Dispersion

وهي مقاييس تعبر عن بعد أو قرب المشاهدات عن بعضها البعض أي عن كيفية انتشارها لتشتت أي تبين مدى انسجام المشاهدات مع بعضها البعض وهي على النحو التالي.

المدى Range

وهو الفرق بين أعلى مشاهدة وأقل مشاهدة ، ويعتبر أكثر المقاييس تطرفا وأقلها ثباتا.

التباين Variance

وهو قيمة تعبر عن الفروق بين المشاهدات ويعرف حسابيا بأنه الوسط الحسابي لمربعات انحرافات المشاهدات عن الوسط الحسابي مقسوما على عدد المشاهدات أو مجموع مربعات انحرافات القيم عن الوسط الحسابي مقسوما على عدد المشاهدات، وكلما زادت قيمته كانت القيم غير متسقة والعكس صحيح، ويمكن حساب التباين لمجموعة من المشاهدات من خلال المعادلة (٢)

$$ع^{٢} = \frac{\text{مجموع (س - س)}^{٢}}{ن} \quad ....(٢).$$

أما بالنسبة للبيانات المبوبة في جداول فيمكن حسابها من نفس المعادلة إنما يتم التعامل مع مراكز الفئات بدلا من المشاهدات ، حيث تتوزع العلامات في فئات، كما يحسب الانحراف المعياري من نفس المعادلة

الانحراف المعياري Standard Deviation

وهو أيضا يعبر عن اتساق المسافات الفروق بين القيم حيث يحسب عادة من قيمة الجذر التربيعي للتباين، حيث تستخدم نفس المعادلة لكن تكون القيمة تحت الجذر التربيعي، وللانحراف المعياري أهمية في التعرف على الخطأ المعياري للتقدير وكذلك في التنبؤ كما سيمر معنا فيما بعد.

الانحراف المتوسط Absolute Mean Deviation

وهو أيضا من المقاييس التي تعتبر مؤشرا على الفروق بين القيم ويعرف حسابيا على انه القيمة المطلقة لمجموع انحرافات القيم عن وسطها الحسابي مقسوما على عددها

ففي المثال السابق يمكن إيجاد قيم التشتت كما في الجدول على النحو التالي:

القيم	٣	٤	٤	٤	٥	٦	٧	٨	٩	١٠	مجموع
الانحراف عن الوسط	٣-	٢-	٢-	٢-	١-	٠	١	٢	٣	٤	صفر
مربع الانحراف	٩	٤	٤	٤	١	٠	١	٤	[illegible]	١٦	٥٢
الانحراف المتوسط	٣	٢	٢	٢	١	٠	١	٢	٣	٤	٢٠

المدى = أعلى قيمة – اقل قيمة = ١٠ – ٣ = ٧

التباين ($ع^{٢}$) مجموع مربعات الانحرافات مقسوما على العدد = ٥٢/ ١٠ = ٥,٢

الانحراف المعياري (ع) = الجذر التربيعي للتباين = ٢,٢٨

الانحراف المتوسط مجموع القيمة المطلقة للانحرافات / العدد ٢٠/ ١٠ = ٢

ملاحظة

هناك حالات معينة يمكن استبدال (ن) الموجودة في المقام ن-١ بهدف التخلص من التحيز الناتج من عن تقدير تباين المجتمع بتباين العينة وفي الحاسبات الآلية يستخدم الرمز (ن) عند استخدامها لحساب التباين للعينة

الإحصاء الاستنتاجي :

ويشير إلى الإحصائيات التي يتم استخراجها من خلال عينة جزئية مختارة من نفس المجتمع خاصة بالمجتمع حيث يتم من خلالها معرفة خصائص المجتمع الكلي ، وتتعلق هذه الإحصائيات بتعميم النتائج على المجتمع شريطة صحة أو توفر بعض الافتراضات منها:

- الاختيار العشوائي لمكونات العينة بحيث تضمن عدم الحصول على نتائج متحيزة لفئة معينة من خصائص المجتمع
- مناسبة الإجراءات الإحصائية المستخدمة لطبيعة البيانات حيث لا تستخدم الإجراءات الإحصائية الواحدة لبيانات مختلفة من حيث قابليتها أو عدم قابليتها للقياس ومستوى القياس الذي تقع عليه ومدى قابلتها للقياس.
- التوزيع الطبيعي للبيانات حيث يعتمد كثير من الأساليب الإحصائية على طبيعة التوزيع الذي تتخذه البيانات إن كان طبيعيا أو طبيعي محول أو معياري . الخ والافتراض الخاص بالتوزيع مفاده أن العينة التي يتم اختيارها من المجتمع يتوزع بشكل طبيعي أي توزيع جرسي أو متماثل او تقاربي يختص بكثافة احتمالية معينة . أما التوزيع الطبيعي المعياري فان الوسط الحسابي الخاص به يساوي (صفر) والانحراف المعياري يساوي (١) وبالرموز فان س= (صفر) وهناك جداول خاصة يعرف منها الاحتمال والذي يساوي المساحة الواقعة تحت المنحنى ضمن مدى معين أو قيم معيارية والتي تعرف عادة بالقيم المفرجة .

وتستخرج العلامة المعيارية من المعادلة رقم (٣)

$$ز_{س} = \frac{س - \bar{س}}{ع_{س}} \quad (٣)$$

حيث:

س : الدرجة الخام (درجة طالب)

$\bar{س}$ - س: الوسط الحسابي لجميع الدرجات (للصف)

ع س. الانحراف المعياري للدرجات جميعها.

ومن المعروف أن الوسط الحسابي للعلامات المعيارية أيضا يساوي(صفر)والانحراف المعياري يساوي (١) شكل (٢).



شكل (٢) التوزيع الطبيعي

معامل الارتباط

توجد في الطبيعة علاقات ارتباطيه بين المتغيرات ومن اهتمامات علم القياس فهم طبيعة هذه العلاقات وهي على نمطين : الأول ارتباط موجب ويسمى أحيانا بالارتباط الطردي مثل العلاقة بين سوء الوضع الاقتصادي للأسرة والحالة الصحية لأفراد الأسرة، حيث يتوقع أن تزداد الحالة الصحية لأفراد الأسرة استقرارا وخلوا من المعاناة من الأمراض كلما كان وضعها الاقتصادي كبيرا . أما النمط الثاني الارتباط السالب ويسمى أحيانا بالارتباط العكسي مثل الارتباط بين العمر والقدرة الجسمية للإنسان حيث تقل القدرة كلما زاد عمر الإنسان ، ومن الجدير ذكره هنا انه ليس

بالضرورة أن يكون هناك ارتباط بين كل المتغيرات حيث لا ترتبط بعض المتغيرات بالبعض الآخر ، ويتم التعبير عن الارتباط أو العلاقة بين متغيرين بقيمة تسمى معامل الارتباط (Correlation- Co-efficient) وتقع قيمته بين القيمتين (-١ - +١) وعندما تكون الإشارة سالبة تكون العلاقة عكسية وإذا كانت موجبة فمعنى ذلك أن العلاقة طردية أي أن الإشارة تحدد نوع الارتباط طردي او سالب ام قوة العلاقة فتبينها القيمة ويظهر ذلك من خلال الجدول (٢) نوع وقوة معاملات الارتباط.

جدول :(٢) نوع وقوة الارتباط بين المتغيرات

القيمة	نوع العلاقة	قوة العلاقة
- ١,٠	عكسية	ارتباط تام
-٠,٧	عكسية	عالي
-٠,٥	عكسية	متوسط
صفر	لا علاقة	لا علاقة
٠,	طردي	ضعيف
٠,٣	طردي	ضعيف
٠,٥	طردي	متوسط
٠,٧	طردي	عالي
١	طردي	تام

ومن المهم أن ننوه هنا إلى أن المتغيرات مختلفة من حيث طبيعتها في قابليتها للقياس وكذلك في مستوى القياس التي تقع عليه ، وتبعا لذلك فهناك عدة صيغ أو معادلات لحساب قيمة معامل الارتباط بيمن المتغيرات حسب طبيعتها وسوف يتم الحديث عن هذه الصيغ على النحو التالي.

دار جرير للنشر والتوزيع

معامل ارتباط بيرسون

ينظر إلى معامل ارتباط بيرسون على انه أكثر معاملات الارتباط انتشارا لأنه عبارة عن الوسط الحسابي لحاصل الضرب العلامات المعيارية للمتغيرين الداخلين في الارتباط ونعبر عنه بالمعادلة رقم (٤) :

$$ر_{س ص} = \sqrt{\frac{\text{مجموع ز س} \times \text{مجموع ز ص}}{ع_{س} \times ع_{ص}}} \quad .... \; .... \; (٤)$$

حيث ز تشير الى العلامة المعيارية أو يمكن التعبير عن ذلك بالمعادلة رقم (٥).

$$ر_{س ص} = \frac{١}{ن} \sqrt{\frac{\text{مجموع ١ س} - \bar{س})(ص - \bar{ص})}{ع_{س} \times ع_{ص}}} \quad ....(٥)$$

حيث :

س: أي درجة خام للمتغير س.

$\bar{س}$: الوسط الحسابي لدرجات المتغير س

ص : أي درجة خام للمتغير ص.

$\bar{ص}$: الوسط الحسابي لدرجات المتغير ص.

$ع_{س}$: الانحراف المعياري لدرجات المتغير س

$ع_{ص}$ الانحراف المعياري لدرجات المتغير ص.

ومن الجدير بالذكر أن حاصل ضرب الجزء الأول (١/ن) في بسط الجزء الثاني من المعادلة يسمى بالتباين المشترك ، حيث انه هو المسئول عن تحديد قيمة معامل الارتباط وهذا يقودنا إلى إمكانية صياغة المعادلة أعلاه إلى الصورة التالية:

$$ر_{س ص} = \frac{ع_{س ص}}{ع_{س} \times ع_{ص}}$$

أي أن التباين المشترك لمتغيرين يساوي الانحراف المعياري للمتغيرين معا مقسوما على حاصل ضرب الانحراف المعياري للمتغير س في الانحراف المعياري للمتغير ص ، ومن هذه الصورة يمكن الوصول إلى الصورة النهائية لمعامل ارتباط بيرسون كما في المعادلة رقم(٦)

$$\frac{\text{ن مجموع س ص} - \text{مجموع س} \times \text{مجموع ص}}{\sqrt{\text{ن }(\text{س}^{2} - (\text{مجموع س})^{2}) \times \text{ن}(\text{ص}^{2} - (\text{مجموع ص})^{2})}} \quad \ldots\ldots(٦)$$

حيث:

ن: عدد المفحوصين.

س: أي درجة للمتغير س.

ص: أي درجة للمتغير ص.

مثال (١)

كانت درجات (١٠) طلاب في اختبار يومي لمادتي العلوم واللغة العربية في احد صفوف المرحلة الإعدادية كما هي في الجدول التالي والمطلوب حساب معامل الارتباط بين هذه الدرجات حسب معامل ارتباط بيرسون.

الرقم	درجة العلوم س	درجة الرياضيات ص	س٢	ص٢	س×ص
١	١٠	١٤	١٠٠	١٩٦	١٤٠
٢	٦	٩	٣٦	٨١	٥٤
٣	٨	١١	٦٤	١٢١	٨٨
٤	٦	١٢	٣٦	١٤٤	٧٢
٥	٩	١٠	٨١	١٠٠	٩٠
٦	٧	١١	٤٩	١٢١	٧٧

٧	٩	١١	٨١	١٢١	٩٩
٨	٥	٨	٢٥	٦٤	٤٠
٩	٦	١١	٣٦	١٢١	٦٦
١٠	٧	١٢	٤٩	١٤٤	٨٤
المجموع	٧٣	١٠٩	٥٥٧	١٢١٣	٨١٠

تحسب القيم في الحقول أعلاه لتوفير مكونات معادلة معامل الارتباط وتظهر في صف المجموع ويكون الحل على النحو التالي

$$\frac{\text{ن مجموع س ص} - \text{مجموع س} \times \text{مجموع ص}}{\sqrt{\text{ن (س}^{٢} - \text{(مجموع س)}^{٢} \times \text{ن(ص}^{٢} - \text{(مجموع ص)}^{٢}}}$$

وبتطبيق المعادلة ينتج

$$\frac{١٠(٨١٠)-(٧٣\times١٠٩)}{\sqrt{١٠(٥٥٧)-(٧٣)^{٢}\times١٠(١٢١٣)-(١٠٩)^{٢}}}$$

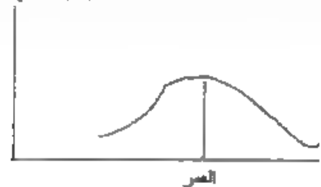
$$\frac{٨١٠٠-٧٩٥٧}{٢٤٤}$$

ر س ص = ٠.١٧

ويوصف معامل الارتباط بدلالة الإشارة والمقدار فالإشارة (+، -) تشير إلى اتجاه العلاقة فهي (أي إشارة) سالبة عندما تكون العلاقة بين المتغيرين عكسية أو سالبة وموجبة عندما تكون العلاقة بين المتغيرين طردية أو موجبة ، وتعني العلاقة الطردية أو الموجبة بين متغيرين أن الأشخاص ذوي العلامات العالية على المتغير(س) يميلون إلى الحصول على علامات عالية على متغير(ص) . أما العلاقة العكسية فتعني أن الأشخاص ذوي العلامات العالية على (س) يميلون إلى الحصول على علامات متدنية على المتغير (ص)

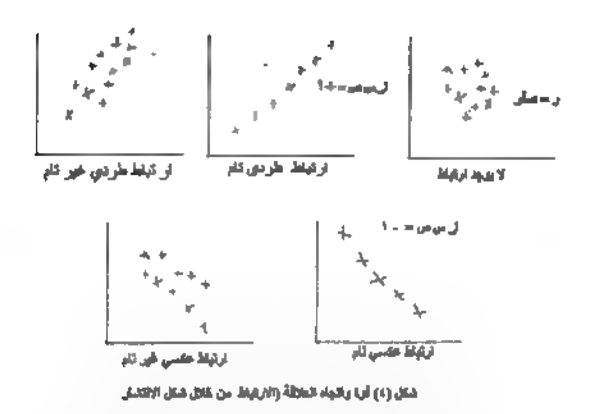
وإذا انتقلنا إلى مقدار معامل الارتباط أو حجمه (بعيدا عن الإشارة) فانه ينحصر بين صفر +١ وهذا يعكس قوة العلاقة أي مدى التنبؤ بقيم أحد المتغيرين من خلال توفر قيم المتغير الثاني وبيان ذلك أنه كلما اقترب حجم معامل الارتباط من العدد (١) كلما ازدادت قوة العلاقة بينهما وهكذا تساوى قوة العلاقة في حال كانت قيم معاملات الارتباط (+٩, ٠) و (-٩, ٠) لكن اتجاه الأول موجبة أي علاقة طردية والثاني سالب أي علاقة عكسية رغم تساويهما في القوة وعندما تكون العلاقة غير تامة فأن إمكانية التنبؤ الدقيقة بأحد المتغيرين من الآخر غير ممكنة إذ تزداد دقة التنبؤ مع ازدياد قيمة معامل الارتباط .ولا بد من أن نتذكر انه إذا كانت قيمة معامل الارتباط (+١،و – ١) فمعنى ذلك أن الارتباط تام.

٣- يقيس معامل الارتباط قوة العلاقة الخطية بين متغيرين كما في الشكل رقم (٣).



شكل (٣) قوة العلاقة الخطية بين متغيرين.

حيث يوضح الشكل طبيعة العلاقة بين الذاكرة الإنسانية مع التقدم في العمر حيث تعرف هذه العلاقة المنحنية (غير خطية) حيث يبين الشكل بأنه رغم أن العلاقة قوية إلا أن قيمة معامل الارتباط تعتبر متدنية بسبب عدم خطية الارتباط ، وهذا يشير إلى معنى أكثر تحديدا لمعامل ارتباط بيرسون على انه قيمة محصورة بين (+١) و (-١) حيث تشير هذه القيمة إلى اتجاه وقوة العلاقة بين متغيرين وتتباين قوة العلاقة بين أي متغيرين حيث يمكن الحكم على قوة واتجاه العلاقة من خلال شكل الانتشار لقيم المتغيرين كما يتضح من الشكل رقم (٤) .



شكل (٤) أنوا واتجاه العلاقة (الارتباط) من خلال شكل الانتشار

ومن الجدير بالذكر أن حجم معامل الارتباط لا يتأثر بأي تحويل خطي لأي من المتغيرين (س، ص) أو لكليهما ، ويبقى التباين المشترك ع س ص = ع س* ص* وتقرأ عين سين ستار ص ستار أي العلاقة بعد التحويل وهذه تعني العلاقة بعد التحويل الخطي ع س* ص* (أ س + ب) (ج ص + د) حيث تعتبر كل من (أ، ب، ج، د) ثوابت ، ولا بد من التذكر بان حجم معامل الارتباط يعتمد على القيمة بغض النظر عن الإشارة حيث تبين الإشارة فقط اتجاه الارتباط طردي أو عكسي فمثلا حجم معامل الارتباط (قوة العلاقة) (ر= +٠,٨) = معامل الارتباط (قوة العلاقة) (ر+ -٠,٨) ،

٧١	٦٤	٧	٨	-١	١
٦٥	٥٥	١٠	١٠	٠	٠
٨٨	٩٠	٢,٥	٢	٠,٥	٠,٢٥
٩٢	٨٨	١	٣	-٢	٤
٧٤	٧٢	٥	٥	٠	٠
المجموع					٢٠,٥

وبتطبيق المعادلة نجد أن ر = ١ – (٦ (٦ ٢٠) / ١٠ (٩٩)) = ١ – ١٢٣ / ٩٩٠ = ٠٫٨٧

تفسير معامل الارتباط :

يفسر معامل الارتباط بدلالة مربعة إذ يشير مربع معامل الارتباط إلى نسبة التباين المفسر في احد المتغير والذي يعزى إلى العلاقة الخطية مع المتغير الآخر فإذا كانت قيمة $ر_{س ص}$ = (٠,٥) فان $ر^{٢}_{س ص}$ (٠,٢٥) وهذا يعني : أن (٠,٧٥) من تباين المتغير (ص) يعزى إلى العلاقة الخطية مع المتغير(س) مع تذكر أن($ر_{س ص}$) تختلف عن($ر_{ص س}$). وجدير بالذكر انه إذا كانت هناك علاقة بين متغيرين (س، ص) مثلا

يكون معامل الانحدار موجبا ، أما إذا كان الارتباط سالبا أي أن العلاقة عكسية تكون قيمة معامل الانحدار سالبة، لاحظ الجزء الأيمن للمعادلة (معامل الانحدار).

$$ر\ س\ ص \times \frac{ع\ ص}{ع\ س}$$

كما ويمكن التنبؤ بقيمة احد المتغيرين (ص) مثلا من قيم المتغير (س) بفترة أي أن قيمة ص قد تكون(قل أو اكبر من قيمة س) اي ان ص = ± س.وهنا نستخدم قيمة ف الحرجة (١.٩٦) والخطأ المعياري للتقدير بالاستفادة من معامل الارتباط بين المتغيرين والانحراف المعياري لقيم المتغير المرغوب بالتنبؤ بقيمته كما سنرى فيما بعد ، قد يتساءل البعض عن كيفية التنبؤ بقيم احد المتغيرين إذا علمت قيم المتغير الآخر ، على الرغم من عدم وجود علاقة ارتباط بينهما؟

الجواب طبعا هنا انه لن يكو بمقدورنا التنبؤ ، وان أفضل تقدير لقيم المراد التنبؤ به هو الوسط الحسابي لذلك المتغير .حيث ستكون قيمة معامل الانحدار مساوية للصفر والسبب في ذلك أن عدم وجود ارتباط بين المتغيرين يعني أن ر = صفر وهذا يؤدي إلى أن يكون القيم التي تضرب بالقيمة صفر مساوية للصفر وتبقى القيمة المضافة وهي الجزء الأخير من المعادلة (+ ص ¯) ومن هنا نقول ان أفضل قيمة تقديرية لأي قيمة للمتغير ص هي الوسط الحسابي لتلك القيم.

مثال:

إذا كانت قيمة معامل الارتباط بين درجات طلبة الصف الخامس الابتدائي في مادتي التربية الرياضية والتربية الفنية (ر= ٠.٧٥) وكانت قيمة (ع س =٢) وكانت (ع ص= ٢.٦) كما كانت قيمة (س ¯ = ٢ ٥) وقيمة (ص¯ = ٣.٨) أوجد قيمة ص إذا كانت قيمة (س= ٦) واحسب قيمة ص بفترة ثقة (٩٥٪).

الحل:

حسب المعادلة:

$$ص^{\wedge} = ر_{س ص} \times \frac{ع_{ص}}{ع_{س}} \times (س - \bar{س}) + (\bar{ص})$$

$= ٠.٧٥ \times (٢.٦/٢) \times (٦ - ٥.٢) + ٣.٨$

$= ٠.٩٧٥ (٠.٨) + ٣.٨$

$= ٠.٧٨ + ٣.٨ = ٤.٥٨$

هذا إذا أردنا التنبؤ بنقطة (بعلامة محددة) أما إذا أردنا التنبؤ بفترة أي أن قيمة ص قد تكون اقل من أو اكبر من قيمة س المعروفة س≥ص≥س فسيكون التقدير كالآتي. وهذا يتطلب حساب الخطأ المعياري للتقدير من المعادلة

$$ع_{ص.س} = ع_{س} \times \sqrt{١ - ر^{٢}_{س.ص}}$$

$= ٢.٦ \times ٠.٦٦ = ١.٧٢$

وهنا ستكون قيمة ص

$=$ س − ف بدرجة × ع ص.س ≤ ص ≤ س + ف بدرجة × ع ص.س

$= ٦ - (١.٩٦ \times ١.٧٢) \leq ص \leq ٦ + (١.٩٦ \times ١.٧٢)$

$= ٦ - ٣.٣٧ \leq ص \leq ٦ + ٣.٣٧$

$= ٢.٦٣ \leq ص \leq ٩.٣٧$

أي ص= س − ٣.٣٧ أو س +٣.٣٧

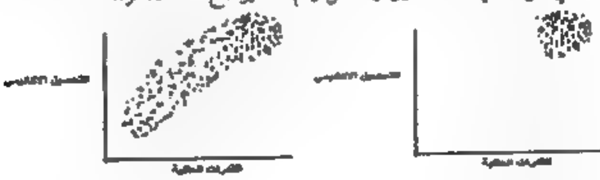
أي انه إذا كانت قيمة س =٦ فان قيمة ص تقع بين (٢.٦٣ − ٩.٣٧) ومعنى مستوى الثقة (٩٥٪) أن (٩٥٪ من المفحوصين الذين تكون درجاتهم على المتغير (س = ٦) أن درجاتهم على المتغير (ص) ستقع ضمن المدى(٢.٦٣ − ٩.٣٧).

العوامل المؤثرة في معامل الارتباط.

يتأثر معامل الارتباط بعدة عوامل ويتمثل هذا التأثير بانخفاض أو ارتفاع قيمة معامل الارتباط ويتعلق التأثير بأحد المتغيرين أو بكليهما ومن أهم العوامل المؤثرة ما يلي

ظاهرة ضيق المدى

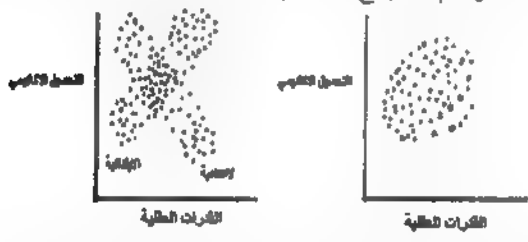
وتتمثل ظاهرة ضيق المدى بوقوع قيم احد المتغيرين أو كليهما ضمن مدى قليل وهنا تميل قيمة معامل الارتباط إلى الانخفاض عن القيمة التي لا تكون فيها قيم المتغيرين واقعة ضمن مدى ضيق ، ويتعلق هذا العامل بأساليب اختيار العينات، وهذا ما يسمى أحيانا بأثر الانخفاض ولتوضيح ذلك دعنا نفترض أن معامل الارتباط بين القدرات العقلية والتحصيل المدرسي (٩٠ ٠) ويتضح هذا الارتباط من خلال شكل الانتشار للدرجات على اختبار القدرات العقلية والتحصيل المدرسي للطلبة كما في الشكل (٥- أ) وبعد دراسة هذا الارتباط لمجموعة من الطلبة ممن قبلوا في إحدى الجامعات والتي يتضح الارتباط بين الدرجات على اختبار للقدرات العقلية والمعدل التراكمي الجامعي كما في الشكل (٥- ب) فانه يتوقع أن يكون معامل الارتباط في الحالة الثانية أقل من الحالة الأولى والسبب في ذلك أن درجات الطلبة على اختبار القدرات العقلية والتحصيل الأكاديمي في الحالة الثانية ستكون واقعة ضمن مدى محدد وأضيق منه في الحالة الأولى ولن يمثل جميع درجات الطلبة الذين تم إدخالهم في الدراسة (الحالة الأولى) بل سيكونون مجموعة جزئية منها (الحالة الثانية) ، حيث تكون الدرجات الأعلى حيث يخفي ضيق المدى هذا جزءا من الارتباط حيث يتوقع أن يكون معامل الارتباط في الحالة الثانية أقل والسبب في ذلك أن قيم المتغيرين ستكون مجموعة جزئية متجانسة نسبيا من المجموعة الكلية الأمر الذي يوحي بان الانحدار الخطي اقل منه في الحالة الأولى والشكل رقم (٥) يوضح هذه الظاهرة.



شكل (٥- أ) درجات تقع ضمن مدى ضيق   شكل (٥- ب) درجات تقع ضمن أكثر اتساعا

دمج المجموعات.

تؤثر عملية دمج المجموعات في أنها قد تخفي جزءا من معامل الارتباط بين متغيرين وتظهر هذه الحالة فيحال وجود ارتباط بين متغيرين لدى مجموعتين أو أكثر فقد يكون الارتباط بين القدرات العقلية والتحصيل الأكاديمي للطلبة ارتباطا ايجابيا ولصالح في المرحلة الابتدائية الأولى بينما قد يكون هذا الارتباط سلبيا في المرحلة الإعدادية ، وهذا يكون إذا درست العلاقة لكل مرحلة على حدة ، أما اذا درست العلاقة على مجموعة من المرحلتين فقد تظهر النتائج غير ذلك بسبب دمج طلبة المرحلتين معا في مجموعة واحدة وهما في الأصل مجموعتين منفصلتين ، فقد تظهر النتائج عن عدم وجود علاقة بسبب اختلاف اتجاه الارتباط بين المتغيرين عند المجموعتين كل على حدة أو أن الارتباط قد يبدو منخفضا، والشكل رقم (٦) يوضح هذه الظاهرة



شكل رقم (٦أ – ٦ ب) ظاهرة دمج المجموعات وأثرها على معامل الارتباط

لاحظ أن العلاقة بين القدرات العقلية والتحصيل الأكاديمي قوية لكنها متعاكسة، الأمر ، حيث أن دمج المرحلتين من لو تمت دراسة العلاقة لكل مرحلة على حدة لاحظ شكل (٦) يسار ، أما إذا أخذنا مجموعة جزئية من المرحلتين (لحظ شكل (٦) يمين فان ذلك سيؤثر هذا الدمج على الارتباط كون كل واحدة جاءت من مجموعة يختلف فيها واقع العلاقة بين المتغيرين ولذلك لا بد للباحثين الاهتمام باختيار العينات لتجنب مثل هذه التأثيرات.

صور أخرى لمعامل الارتباط

إن معامل ارتباط بيرسون يتطلب توفر بعض الافتراضات أو الخصائص في المتغيرات ، مثل مستوى القياس الذي تقع عليه والذي يتحدد من خلال طبيعة المتغير كأن يكون متصلا أو منفصلا أو محولا ..الخ .فمعامل ارتباط بيرسون مثلا يستخدم لتحديد العلاقة بين متغيرين كلاهما يقع على مستوى قياس فئوي أو اعلي أي متغير متصل السؤال الذي يطرح نفسه الآن هو هل يصلح معامل ارتباط بيرسون لحساب العلاقة بين المتغيرات المنفصلة مثل الجنس(ذكر ، أنثى) والدرجة على اختبار للتحصيل فقراته من نوع الإجابة المنتقاة ببديلين او اكثر أو أن الدرجة عليه تأخذ (صفر، ١) مثلا، في الواقع إن طبيعة المتغيرات تؤثر في حسب معامل الارتباط فقد يكون احد المتغيرين متصلا والأخر منفصلا يتم تحويله كي يهل التعامل معه ، ومن هنا جاء تعدد صيغ معامل الارتباط لتكيف مع اختلاف طبيعة المتغيرات. ومن الصيغ الأخرى لمعامل الارتباط ما يلي :

معامل الارتباط فاي (Φ):

يعرف هذا المعامل بأنة معامل ارتباط بين متغيرين كل منهما منفصل ثنائي بصورة طبيعية ومعنى منفصل أن المتغير يقع على مقياس اسمي مثل (الجنس ، الإجابة عن سؤال من بديلين) فإذا اردنا أنتكشف عهن وجود ارتباط بين الجنس والإجابة على فقرات اختبار من نوع (نعم، لا أو صح ، خطأ) وإيجاد قيمة هذا الارتباط فان الصيغة المناسبة لذلك هي معامل ارتباط فاي (Φ) من خلال المعادلة رقم (١١).

$$\text{ر فاي} = \frac{\text{ح س ص} - \text{ح س ح ص}}{\sqrt{\text{ح س} \times (1 - \text{ح س}) \times (1 - \text{ح ص}) \times \text{ح ص}}} \quad \ldots\ldots(١١)$$

حيث

ح س ص : نسبة الأفراد ذوي العلامة ١ على المتغيرين

ح س : نسبة " " ١ " للمتغير س

ح ص : " " " ١ " " " ص

أ ـــ ح س: صفر س وهي متممة نسبة المفحوصين الذين حصلوا على (١) على س.

أ ـــ ح ص: صفر ص وهي متممة نسبة المفحوصين الذين حصلوا (١) على ص.

تفسير معامل الارتباط ، Φ فاي

يفسر معامل الارتباط فاي تفسير معامل ارتباط بيرسون حيث تشير الإشارة (سالب، موجب) إلى اتجاه العلاقة وتشير القيمة إلى قوة العلاقة، ويعتبر معامل ارتباط فاي حالة خاصة من معامل الارتباط بيرسون والفرق بين المعاملين هو انه في معامل بيرسون نستخدم الدرجات الخام ، أما في معامل فاي فإننا مضطرون إلى تصنيف المفحوصين حسب طبيعة إجاباتهم وبالتالي حسب طبيعة درجاتهم التي حصلوا عليها، أي إننا في معامل فاي سنستخدم النسب المئوية لأصناف الإجابات ، وهذا يعني أن معامل ارتباط فاي هو حالة خاصة من معامل ارتباط بيرسون ، ولذلك فان الارتباط التام يمكن الحصول عليه وفقا لمعامل بيرسون ، أما حسب لمعامل فاي فان الارتباط التام مرهون بتوفر ظرف وهو أن تكون نسبة من يجيبوا على الفقرتين بنعم مساوية للقيمة (٠.٥) .

وبملاحظة كل من معاملي الارتباط (بيرسون و فاي Φ) نجد إنهما متشابهين ويختلفان بفارق بسيط وهو استخدام العلامة الخام في معامل بيرسون والنسبة في معامل فاي ارتباط ويمكننا القول أن معامل ارتباط فاي هو حالة خاصة من معامل ارتباط بيرسون ويتوقع أن تكون قيمته النظرية محصورة بين (+١و- ١) ومن الصعب عمليا الحصول على ارتباط تام إلا في حال تساوت نسبتي من أجابوا نعم على المتغيرين س و ص. وكذلك ، إلا في حالة واحدة وهي عندنا وتكون × أ = أ أي أن نسبة نجاح المتغير الأول = نسبة نجاح المتغير الثاني وتساوي (٠,٥).

مثال :

تقدم (٤٠) طالب إلى اختبار تحصيلي في اللغة العربية من نوع الاختيار من متعدد، سحبت إجابات الطلبة على فقرتين من الاختبار وكانت كما تظهر في الجدول التالي أوجد قيمة معامل ارتباط (Φ) بين الفقرتين من خلال إجابات الطلبة مع بيان نوع الارتباط وكذلك تفسير معامل الارتباط

الفقرة س	الإجابات		المجموع
الفقرة ص	لا	نعم	
نعم	١٢ (ج)	١٦ (أ)	٢٨
لا	٨ (ب)	٤ (د)	١٢
المجموع	٢٠	٢٠	٤٠

الحل : اعتماد النسب حسب المعادلة التالية :

ح س ص = ١٦/ ٤٠ = ٠.٤

ح س = ١٢+ ١٦/ ٤٠ = ٢٨/ ٤٠ = ٠.٧

ح ص : ١٦+ ٤/ ٤٠ = ٢٠/ ٤٠ = ٠.٥.

١_ح س: ١ − ٠.٧ = ٠.٣.

١_ح ص . ١ − ٠.٥ = ٠.٥.

$$\text{ر} = \frac{\text{ح س ص} - \text{ح س ح ص}}{\sqrt{\text{ح س} \times (1 - \text{ح س}) \, \text{ح ص} \times (1 - \text{ح ص}) \times \text{ح ص}}}$$

$$\text{ر} = \frac{0.4 - (0.7 \times 0.5)}{\sqrt{0.3 \times 0.7) \, (0.5 \times 0.5)}}$$

$$\text{ر} = \frac{0.05}{\sqrt{0.21 \times 0.25}}$$

$$ر_{فاي} = \frac{0.05}{\sqrt{0.0525}}$$

$= 0.05 / 0.23 = 0.22$

وتفسير ذلك أن الارتباط بين المتغيرين(س، ص) ارتباط موجب ومتدني نوعا ما.

كما يمكن حساب معمل الارتباط باستخدام عدد المفحوصين بعد توزيعهم على احتمالات الإجابة لكلا المتغيرين من خلال المعادلة رقم (11 – ب).

$$ر_{فاي} = \frac{(أ \times ب) - (جـ \times د)}{\sqrt{(أ+جـ) \times (ب+د) \times (أ+د) \times (ج+ب)}} \quad ....(11 - ب)$$

ولتوضيح ذلك سنعود للجدول في المثال السابق وسيرمز إلى فئات المفحوصين حسب إجاباتهم كما هو واضح في الجدول اعلاه على النحو التالي:

أ: عدد المفحوصين الذين أجابوا على (س ، ص) إجابة صحيحة.

ب: عدد المفحوصين الذين أجابوا على (س، ص) إجابة خاطئة.

ج: عدد المفحوصين اللذين أجابوا على (س) إجابة صحيحة وعلى(ص) إجابة خاطئة.

د: عدد المفحوصين الذين أجابوا على (س) إجابة خاطئة وعلى (ص) إجابة صحيحة. وبتطبيق المعادلة ينتج

$$ر_{فاي} = \frac{(16 \times 8) - (12 \times 4)}{\sqrt{(16+12)\times(8+4)\times(12+8)\times(16+4)}}$$

$$ر_{فاي} = \frac{128 - 48}{\sqrt{(28) \times (12) \times (20) \times (20)}}$$

ر فاي $80 / 366.6 = 0.219$

معامل الارتباط الرباعي (Tetra choric)

كنا تحدثنا سابقا أن طبيعة المتغيرات مختلفة والكشف عن العلاقة بينها يتطلب إجراءات تتناسب مع طبيعة هذه المتغيرات فمعامل ارتباط بيرسون يبين الارتباط بين متغيرين كل منهما متصل ، ومعامل فاي يبين الارتباط بين متغيرين كل منهما ثنائي الفئة ومنفصل بالطبيعة بحيث تأخذ كل فئة قيمة معينة أي انه مكون من فئتين أما إذا كان المتغيرين منفصلين وتم تحويلهما على افتراض أنهما يتوزعان توزيعا طبيعيا وأنهما متصلان من خلال عملية التحويل وأن العلاقة بينهما علاقة خطية وهنا فان صيغة معامل الارتباط بينهما مختلفة ويعرف معامل الارتباط بمعامل ارتباط تراشورك (Tetrachoric Coefficient) ويسمى أحيانا بمعامل الارتباط الرباعي ويتم إيجاد قيمة معامل ارتباط تراشورك من المعادلة رقم (١٢)

$$\frac{(\text{أ} \times \text{ب}) - (\text{جـ} \times \text{د})}{\text{ن}^{2}\ (\text{ف س} \times \text{ف ص})} = \text{ر} + \frac{\text{ب س} \times \text{ب ص ر}^{2}}{2} \quad \text{..........................} (١٢)$$

وتتضح الرموز أ، ب ، جـ دل) من خلال الجدول التالي

أما (ف س ، ف ص) فهي قيمة تستخرج من جداول مساحة المنحنى الاعتدالي المعياري ثم إيجادها من خلال المساحة (الارتفاع) المقابل للفرق بين النسب المكونة للخلايا العمودية والأفقية في الجدول أعلاه،على النحو التالي:

	س		المجموع
ص	أ	ب	أ + ب %
	جـ	د	جـ + د %
	أ + جـ %	ب + د %	الكلي

ف س : هي الارتفاع الفصل بين النسبتين (أ+ ب) و (جـ + د).

ف ص: الارتفاع الفاصل بين النسبتين (أ+ جـ) و (ب + د).

ب س : بعد المتوسط عن الارتفاع ف س بالدرجات المعيارية

ب ص بعد المتوسط عن الارتفاع ف ص بالدرجات المعيارية.

وهذا يفترض بالطبع بأن كلا المتغيرين يتوزعان أو يتخذان شكل التوزيع الاعتدالي، وهنا يتم تكوين جدول كما هو مبين أعلاه حيث يعتمد على عدد فئات المتغيران ولحساب قيمة معامل الارتباط بهذه الطريقة يتم استخدام جداول خاصة تتضمن قيم المساحة الخاصة بالنسب تحت المنحنى الطبيعي وتسمى بالجدول الرباعي لنسب المقاييس الثنائية حيث يتم حساب النسب ضمن كل خلية من خلايا الجداول والتي تشير إلى توزيع الأفراد عينة الارتباط على فئات المتغيرين معا ولتوضيح كيفية حساب قيمة معامل الارتباط دعنا نأخذ المثال التالي.

مثال :

أرادت إحدى المؤسسات دراسة العلاقة بين مستوى التحصيل الأكاديمي لمنتسبيها وتكيفهم مع متطلبات العمل في المؤسسة من خلال تصنيفهم إلى ذوي مستوى عالي ومتوسط من خلال اختبار تحصيل عام وكذلك تصنيفهم إلى متكيفين وغير متكيفين من خلال اختبار يقيس التكيف مع متطلبات العمل حسب الجدول التالي والمطلوب حساب معامل الارتباط بين التحصيل والتكيف

تحصيل / تكيف	عالي	متوسط	المجموع
متكيف	٣٥ (أ)	٢٥ (ب)	٦٠
غير متكيف	١٠ (جـ)	٣٠ (د)	٤٠
المجموع	٤٥	٥٥	١٠٠

دار جرير للنشر والتوزيع

الحل:

لتطبيق المعادلة لا بد من توفير القيم المكونة لها ، ومن خلال الجدول تتبين القيم الخاصة بالرموز(أ، ب، جـ د) ، وأما القيم الخاصة فهي على النحو التالي:

أ عدد العاملين من المتكيفين ومستوى التحصيل العالي. وعددهم (٣٥)

ب. عدد العاملين من المتكيفين ومستوى التحصيل المتوسط. وعددهم (٢٥).

ج: عدد العاملين من غير المتكيفين ومستوى التحصيل العالي. وعددهم (١٠).

د: عدد عدد العاملين من غير المتكيفين ومستوى التحصيل المتوسط. وعددهم (٣٠).

ف س هي الارتفاع المقابل للفرق بين النسبتين (أ + ب) و (جـ + د) وهما (٦٠٪) و (٤٠٪)وهي (٣٨٦ ٠) من الجدول الاعتدالي

ف ص. الارتفاع المقابل للفرق بين نسبتي (أ+جـ) و (ب + د) وهما (٤٥٪) و (٥٥٪). وهي (٠.٣٩٦) من الجدول الاعتدالي

ب س : بعد المتوسط عن الارتفاع ف س بالدرجات المعيارية. وتساوي (-٠.٢٥٣)

ب ص: بعد المتوسط عن الارتفاع ف ص بالدرجات المعيارية. وتساوي (٠.١٢٦)

وبتطبيق المعادلة (١٢) ينتج أن :

$$\frac{(\text{أ} \times \text{ب}) - (\text{جـ} \times \text{د})}{\text{ن}^2 (\text{ف س} \times \text{ف ص})} = \text{ر} + \frac{\text{ب س} \times \text{ب ص}\ \text{ر}^2}{2} \quad \ldots\ldots\ldots (12)$$

$$\text{ر ت} = \frac{(35) \times (25) - (25) \times (10)}{(100)^2 \times 0.386 \times 0.396} = \text{ر} + \frac{(-0.253) \times (0.126)\ \text{ر}^2}{2}$$

وبحل المعادلتين فان رتت= (٠.٥٣).

معامل الارتباط ثنائي التسلسل(Point Correlation biserial Coefficient)

وهو معامل بين متغيرين بحيث يكون احد هما ثنائيا منفصلا بصورة طبيعية مثل (الجنس ،نوع الإجابة عن سؤال من نوع الاختيار من بديلين) الثاني متغير يقع على مستوى قياس فئوي أو نسبي مثل (كرواتب موظفين أو درجات المفحوصين على اختبار تحصيلي) ويتم إيجاده من المعادلة رقم (١٣) ·

$$\text{ع س } (\text{١} - \text{ح}) = \frac{\bar{\text{س}}_{\text{١}} - \bar{\text{س}}}{\text{ع س}} \sqrt{\frac{\text{ح ص}}{(\text{١} - \text{ح}) \times \text{ص}}} \quad \ldots \quad (\text{١٣})$$

حيث.

س : درجات المتغير الثنائي ورمزه ص.

ص: درجات المتغير المتصل.

$\bar{\text{س}}$: الوسط الحسابي لدرجات المتغير المتصل (س)

$\bar{\text{س}}_{\text{١}}$: الوسط الحسابي لدرجات المفحوصين على المتغير (ص) ذوي الدرجة (١) على المتغير (س)

ع س: الانحراف المعياري لدرجات المفحوصين على المتغير (س)

ح ص. نسبة المفحوصين الحاصلين على الدرجة (١) على المتغير (ص).

١ – ح ص: نسبة المفحوصين الحاصلين على الدرجة (صفر) على المتغير (ص)

مثال:

خضع (٢٠) مفحوصا لاختبار في الرياضيات من نوع الاختيار من متعدد مكون من ٣٠ فقرة ، فإذا كانت الفقرة الأولى تقيس حل المعادلة بمجهول واحد ، فما هو نوع

الارتباط بين الإجابة على هذه الفقرة والإجابة على فقرات الاختبار ككل، إذا كانت الدرجات كما في الجدول الجانبي.

الرقم	د ف١	د الكلي	الرقم	د ف١	د الكلي	الرقم	د ف١	د الكلي	الرقم	د ف١	د الكلي
١	صفر	١	٦	صفر	٤	١١	١	٨	١٦	١	٥
٢	صفر	٤	٧	١	٥	١٢	١	٩	١٧	صفر	٦
٣	صفر	٦	٨	١	٧	١٣	١	٧	١٨	١	٨
٤	صفر	٥	٩	صفر	٩	١٤	١	٤	١٩	صفر	٩
٥	صفر	٧	١٠	١	٦	١٥	صفر	٥	٢٠	١	٤
الكلي	صفر	٢٣		٣	٣١		٤	٣٣		٣	٣٢

س ‾‾:الوسط الحسابي لدرجات المتغير المتصل (س)

١+٤+٦+٥+٧+٤+٥+٧+٩+٦+٨+٩+٧+٤+٥+٥+٦+٨+٩+٤ /٢٠= ١٢٠/٢٠ = ٦

س١‾‾ : الوسط الحسابي لدرجات المفحوصين ذوي الدرجة (١)على المتغير (ص)

وهي هنا تمثل الوسط الحسابي لدرجات الطلبة الذين أجابوا الفقرة إجابة صحيحة وبملاحظة الجدول أعلاه نلاحظ أنهم ذوي الأرقام(٧، ٨، ١٠، ١١، ١٢، ١٣، ١٤، ١٦، ١٨، ٢٠)

٥+٧+٦+٨+٩+٧+٤+٥+٨+٤/ ١٠= ٦٣/ ١٠= ٦.٣

ع س:الانحراف المعياري لدرجات المفحوصين على المتغير (س)

وبحسابه نجد التباين ومن ثم نأخذ قيمة الانحراف المعياري المتمثلة بالجذر التربيعي للتباين من خلالا المعادلة التالية

مجموع (س ـــ س ‾‾)٢ / ٥

مج(-٥^٢ +-٢^٢ +٠+-١^٢ +١^٢ +-٢^٢ +-١^٢ +١^٢ +٣^٢ +٠+٢^٢ +-٣^٢ +١^٢ +-٢^٢ +-١^٢ + -١^٢ +٠+٢^٢ +٣^٢ +-٢^٢)

ع س = ٨٢ / ٢٠ = ٤.١ وبأخذ قيمة الجذر لها فان الانحراف المعياري = ٢.٠٢

ح س: نسبة المفحوصين الحاصلين على الدرجة (١) على المتغير (ص) وهي هنا الفقرة الأولى.

ولحسابها نحسب عدد من أجابوا على الفقرة الأولى إجابة صحيحة ونقسم على العدد الكلي للمفحوصين.

عددهم ١٠ إذن ح ص ١٠/ ٢٠ = ٠.٥

١ – ح ص: نسبة المفحوصين الحاصلين على الدرجة (صفر) على المتغير (ص) وهي هنا الفقرة الأولى

= ١ ــ ٥ . ٠ = ٠.٥ وقيمة الجذر التربيعي لها = ٠.٧٠٧

والآن نطبق المعادلة رقم ١٣

$$\frac{\text{س}_1^- - \text{س}^-}{\text{ع س}} \sqrt{\frac{\text{ح ص}}{(1-\text{ح})\times\text{ص}}} \quad .... (13)$$

$$\frac{(٦.٣_٦)}{٢.٠٢} \sqrt{\frac{٠.٥}{٠.٥}} \quad .... (13)$$

$$\frac{(٦.٣_٦)}{٢.٠٢} \sqrt{\frac{٠.٧٠٧}{٠.٥}}$$

٠.١٤ × ١٠.١٤ = ٠.٢٠٩

وهي قيمة مساوية تقريبا للقيمة في المثال السابق

معامل الارتباط بايسيريال.(bisorial Coefficient Correlation)

قلنا أن تعدد طبيعة المتغيرات وتعدد المواقف المختلفة فرضت تعدد صيغ معامل الارتباط ويأتي معامل الارتباط بايسيريال تلبية لهذا التنوع وهو معامل يحسب العلاقة بين متغيرين احدهما يقع على مقياس فتري أو رتبي أو نسبي والآخر متغير ثنائي منفصل ولكن بصورة غير طبيعية أي انه محول أي انه كان متغير متصلا وموزعا توزيعا اعتداليا تم تقسيم المشاهدات أو الدرجات إلى فئتين من خلال علامة

عددة كأن تكون علامة قطع مثل (متمكن، غير متمكن) أو أن المتغير لا يحتمل إلى فئتين في الأصل مثل (ذكر ، أنثى) بالنسبة للمتغير الآخر، أما المعادلة التي يمكننا حساب قيمة هذه المعامل من خلال المعادلة رقم (١٤)

$$\frac{\overline{س_1} - \overline{س}}{ع_س} \times \frac{ح_ص}{ب_س} \quad ....(١٤)$$

س درجات المتغير الثاني ورمزه س

ص: درجات المتغير المتصل.

$\overline{س}$: الوسط الحسابي لدرجات المتغير المتصل (س)

$\overline{س_1}$: الوسط الحسابي لدرجات المفحوصين على المتغير (ص) ذوي الدرجة (١) على المتغير (س).

ع س. الانحراف المعياري لدرجات المفحوصين على المتغير (س).

ح ص: نسبة المفحوصين الحاصلين على الدرجة (١) على المتغير (ص)

ب س. بعد المتوسط عن الارتفاع المقابل ل ح س بالدرجات المعيارية.

ولتوضيح كيفية حساب قيمة معامل الارتباط بايسيريال يمكن أن نعود للمثال السابق حيث القيم نفسها باستثناء بعد المتوسط عن الارتفاع ب س حيث أن قيم مكونات المعادلة (١٤) هي كما المثال السابق على النحو التالي:

$\overline{س}$ = ٦ ، $\overline{س_1}$ ٦.٣ ، ع س ٢.٠٢ ، ح ص ٠.٥ ، ب س ٠.٣٥٢١ = ٠.٣٥

وبتطبيق المعادلة فان

ر بسيريال = (٦.٣ ـ ٦)/(٢.٠٢)x(٠.٥/ ٠.٣٥)

ر بسيريال = ٠.١٥x ١.٤٣= ٠.٢١

ومن خلال ما سبق يتبين أن لكل معامل ارتباط صيغة تناسب المواقف المختلفة، وكذلك طبيعة المتغيرات من حيث وجودها في الطبيعة ، وقبولها للقيمة الرقمية أو

الرتبة أي مستوى القياس الذي تقع عليه، ومتى وعي العاملون والمهتمون بالقياس هذه الحقيقة فإن ذلك يؤدي إلى مزيد من الدقة في إجراءاتهم ذات العلاقة بالقياس وستتضح معالم حقل القياس في الفصول التالية.

إن الهدف من التقديم للمفاهيم الأساسية المتعلقة بالقياس وكذلك المفاهيم الإحصائية الأولية أو ما يسمى بالإحصاء الوصفي وذلك تسهيلا على الدارس وكذلك القارئ تمهيداً لموضوعات هذا الكتاب

الباب الثاني

النظرية التقليدية في القياس

الفصل الأول

النظرية التقليدية في القياس

مقدمة

يشير مفهوم القياس في علم النفس إلى كيفية استخدام الأرقام ولماذا تستخدم في هذا الميدان من المعرفة ؟ ويتمثل الهدف الأساسي للنظرية في علم القياس سواء حسب النظرية التقليدية أو النظرية الحديثة في تحديد العلاقة بين استجابات الأفراد على اختبار معين والسمة الكامنة وراء هذه الاستجابات ، والأمر الأكثر أهمية في القياس وخاصة القياس النفسي والتربوي بشكل عام هو تحديد مقدار السمات الكامنة وراء أداء الأفراد على الاختبارات المختلفة والاستفادة منها في تفسير النتائج والتنبؤ بسلوكهم في مواقف مماثلة ، واتخاذ قرارات معينة بشأنهم في ضوء هذا التقدير الكمي للسمات.ولذلك فان النظريتين متوافقتين من حيث الهدف ، لكن الاختلاف قد يكمن في الإجراءات والاعتبارات التي تنطلق منها كل من النظريتين

ولكل نظرية من النظريات مجموعة من الفروض والمسلمات تقوم عليها من أجل تفسير الظواهر التي ترتبط بها ، ولابد أن تكون لهذه النظرية القدرة على التفسير والتحليل حتى تكون نظرية صالحة للاستخدام والتطبيق ، وبالنسبة للنظرية التقليدية في القياس فإنها تستند إلى على أربع مسلمات

- أداء الفرد يمكن قياسه وتقديره.
- أداء الفرد إنما هو دالة لخصائصه.
- الخاصية والأداء والعلاقة بينهما تختلف من فرد لآخر الفروق الفردية.
- القياس الظاهري الكلي يتكون من قياس حقيقي وآخر يرجع إلى الخطأ

وتهتم هذه النظرية بالبحث عن الدرجة الحقيقية للفرد من خلال مجال محدد ، على افتراض أن درجات الخطأ للأفراد تكون عشوائية وغير مرتبطة ببعضها ، وذلك

لتطبيقات متوازية للاختبار ويكون متوسط درجات الخطأ هذه مساوياً للصفر ، كما أن درجات الخطأ تكون غير مرتبطة بالدرجات الحقيقية ، وأن درجات الخطأ والدرجات الحقيقية والدرجات الملاحظة تكون مرتبطة خطياً ويعبر عنها بنموذج الدرجة الحقيقية (True Score)الذي يأخذ شكل المعادلة ك = ح + خ .

حيث

ك: الدرجة الملاحظة

ح: هي الدرجة الحقيقة

خ: درجة الخطأ.

وقد سيطرت نظرية القياس التقليدية على حركة القياس فترة ليست بالقصيرة ، فقد استخدمت أسس هذه النظرية في مواقف اختباريه مختلفة تتضمن بناء مختلف أنواع الاختبارات النفسية ، وكذلك تحليل البيانات المستمدة من هذه الاختبارات ،معتمادا على الافتراضات الخاصة بها.وبالرغم من سيطرة وانتشار تطبيق هذه النظرية وما أرتبط بها من مقاييس إحصائية خاصة بتحليل مفردات الاختبار، إلا أنه تبين قصور هذه النظرية في مواجهة كثير من المشكلات السيكومترية المعاصرة ولمتعرف على مضامين هذه النظرية لا بد من التعرف على المفاهيم الخاصة بها ومن ثم التعرف على إجراءاتها الخاصة بعملية القياس، حيث ان هناك العديد من الاستخدامات والتحليلات والتفسيرات المعتمدة على طبيعة هذه النظرية.

ويرى "جريجوري" (Gregory ,2004) أن نظرية القياس التقليدية بدأت من فكرة أن درجة الفرد على الاختبار تنتج من تأثير مجموعتين من العوامل هما : عوامل تؤدى إلى الاتساق ، وهي عوامل مرغوبة وتتكون من صفات مستقرة لدى الفرد يتم قياسها من خلال الاختبار • والثانية عوامل تؤدى إلى عدم الاتساق أو إلى الاختلاف وتتضمن مجموعة عوامل غير مرغوبة تؤثر في درجة الفرد على الاختبار وتسمى بعوامل الخطأ ، ولغرض تفسير الدرجة من حيث صحتها من عدمه تضع هذه النظرية مجموعة من الافتراضات حددها كل من هامبلتون وزال (Hambleton & Zaal, 1991) على النحو التالي:

- الدرجة الحقيقية للفرد يفترض أن تكون درجة أو قيمة ثابتة ذلك أنها تمثل قدرة الفرد المقاسة.
- الدرجة التي يحصل عليها الفرد ليس من الضروري أن تمثل درجته الحقيقية ولذا فالدرجة التي يحصل عليها قابلة للتغير حسب الظروف الاختبارية .
- الدرجة التي يحصل عليها الفرد هي نتاج نوعين من الدرجات درجة حقيقية ودرجة الخطأ .
- الدرجة الخطأ ليست محددة وثابتة في كل المواقف والظروف بل تتغير بتغير هذه المواقف والظروف.
- وجود علاقة عكسية بين الدرجة الخطأ والدرجة الحقيقية ، وهذا يعني أن انخفاض خطأ القياس يترتب عليه زيادة الدرجة الحقيقية.
- إن الدرجة الحقيقية يمكن معرفتها من خلال تكرار تطبيق الاختبار واستنتاج متوسط الدرجات لهذه التكرارات.
- عدم وجود اقتران بين الدرجات التي يحققها الأفراد وبين الدرجات الخطأ.
- عدم وجود ارتباط بين الدرجات الخطأ في الاختبارات المختلفة وهذا يرجع إلى الاختلاف في طبيعة الاختبارات.
- درجات الخطأ ليست منتظمة ، أي أنها لا تتكرر بنفس الصورة وبنفس المستوى في كل الحالات التي يتم بها تطبيق الاختبار.
- أن الدرجات التي يحصل عليها الفرد في فقرات الاختبار يمكن جمعها كما لو كانت تمثل ميزاناً خطياً Linear Scale ، وأن الفقرات المتعلقة بالمتغير المراد قياسه تحمل المعنى نفسه لدى جميع المفحوصين

وعلى الرغم من استخدام الباحثين للنظرية التقليدية في بناء الاختبارات وتحليل وتفسير نتائجها ، إلا أن هناك بعض المشكلات التي تقلل من دقة وموضوعية هذا الاستخدام ، وهذه المشكلات هي :

(١) عدم وجود وحدة قياس ثابتة : حيث لا تحدد مواضع القياس على متصل المتغير بصورة خطية ، فاعتماد درجات الأفراد على مفردات الاختبار قد يؤدي إلى اختلاف المسافة بين كل درجتين متتاليتين ، ويؤدي هذا إلى اختلاف المعنى الكمي لأي فرق محدد عبر مدى درجات الاختبار(كاظم ، ١٩٩٦).

(٢) تأثر خصائص فقرات الاختبار بقدرة الأفراد : حيث تختلف معاملات الصعوبة أو السهولة والتمييز لفقرات الاختبار باختلاف قدرة أفراد العينة ، فالفقرة التي يختبر بها أفراد ذوي قدرات عالية تبدو سهلة ، بينما تبدو نفس الفقرة صعبة لذوي القدرات المنخفضة. وإذا كانت العينة متجانسة نسبياً ، فإن قيم معاملات التمييز تكون أقل من القيم التي تحصل عليها من عينة غير متجانسة(Hambleton & Swaminathan , 1989).

(٣) تأثر الدرجة الكلية للفرد في اختبار ما بفقراته . حيث تكون درجة الفرد عندما يختبر بفقرات سهلة أعلى منها في حال الفقرات الصعبة ، فلا يمكن تقدير قدرته فيما تقيسه هذه الفقرات تقديراً دقيقاً ، لذا تختلف نتيجة القياس باختلاف الاختبار المستخدم.

(٤) تقتصر الموازنة بين الأفراد في السمة أو القدرة التي يقيسها الاختبار على تطبيق نفس فقرات الاختبار أو مجموعة فقرات مكافئة أو موازية لها على كل فرد من الأفراد. وبالتالي لا نستطيع الموازنة بين مستويات القدرة إذا أجاب الأفراد على مفردات مختلفة ومتباينة في صعوبتها(عماد عبد المسيح ، ١٩٩١).

(٥) تأثر ثبات الاختبار بالموقف الاختباري : حيث يعتمد ثبات الاختبار في إطار هذه النظرية إما على تطبيق الصورة الاختبارية مرتين على أفراد العينة ، أو على إعداد صور متكافئة من الاختبار ويعد هذا في الواقع أمراً صعباً ، وبالرغم من أهمية ذلك ، إلا أنه غير كاف ، حيث يمكن أن يختلف الموقف الاختباري وظروف التطبيق في هاتين المرتين ، حيث اعتبر كل من هامبلتون وسواميثان(Hambleton & Swaminathan , 1989:5) أن هذا الأمر الذي يؤثر على دقة ثبات الاختبار

(٦) تساوى تباين أخطاء القياس لجميع أفراد العينة موضع الاختبار ، وهذا بالرغم من أنه قد يكون أداء بعض الأفراد على الاختبار أكثر اتساقاً من غيرهم من الأفراد وأن درجة هذا الاتساق تختلف باختلاف مستوى قدرة الأفراد أو مستوى القدرة التي يقيسها الاختبار(Randall,1998:6)

(٧) لا تقدم النظرية التقليدية تفسيراً سيكولوجياً يوضح كيف يحاول الفرد إجابة إحدى مفردات الاختبار ، على الرغم من أن هذا التفسير يعد ضرورياً ولازماً إذا أردنا التنبؤ بخصائص الدرجات المستمدة من مجتمع معين أو مجتمعات مختلفة من الأفراد ، أو إذا أردنا تصميم اختبارات تتميز بخصائص سيكومترية معينة تناسب مجتمعاً من الأفراد هذا بالإضافة إلى أن تكوين مفردات الاختبار ومعناها تتغير بتغير عامل الزمن ، أي بمضي الزمن بالنسبة لعينة الأفراد الذين أعد لهم الاختبار ، فالظروف البيئية تتغير ، والظروف الاختبارية ليست دائماً مقننة كما أن حذف أو تغيير أي مفردة من مفردات الاختبار يؤدى إلى تغيير في درجات الأفراد ، هذا التغيير يصعب التنبؤ به (علام ، ١٩٨٥) .

(٨) جميع خصائص الاختبارات التي تستند في بنائها على أسس النظرية التقليدية ، مثل معاملات الصعوبة والتميز والثبات ، تعتمد على خصائص عينة الأفراد التي يجرى عليها الاختبار ، وعلى خصائص عينة الفقرات التي يتكون منها الاختبار.

وفي ضوء ما سبق ، وحيث أن النظرية التقليدية للقياس تؤدى إلى بناء اختبارات غير مرنة ، فقد وجه المتخصصون في القياس جهودهم لوجود نظام قياس أكثر موضوعية يركز على انتقاء المفردات الاختبارية بشكل افضل ، ويسمح بإضافة أو حذف مفردات إلى الاختبار دون أن يتأثر الاختبار ككل ، وقد أدت الجهود إلى ظهور ما يطلق عليه نماذج السمات الكامنة .

وحيث أن النشاط الإنساني في هذا الكون لا يخلو من الممارسات التي تتطلب توفير المعلومات بشقيها الكمي والنوعي حول السلوك والنشاط الإنساني لفهم طبيعة الحياة ومكوناتها وذلك لتسخير هذه المعلومات لفهم هذا الكون بمكوناته وتفصيلاته

وأحداثه والأنشطة التي تتم فيه، ومن ثم توظيفها لخدمة الإنسانية أو لفهم كيفية سير الكون بمكوناته المختلفة ، ولا يمكن توفير المعلومات إلا من خلال علم القياس والذي بدأ ينظم إجراءاته منذ زمن بعيد خلال نظريات القياس بدءاً بنظرية القياس الكلاسيكية أو التقليدية (Classical Test Theory) وانتهاء بنظرية القياس الحديثة والتي عرفت بنظرية السمات الكامنة (latent Trait Theory) حيث تستند كل واحدة من هاتين النظريتين على مجموعة من الافتراضات حول طبيعة السمات الإنسانية وأدوات القياس وخصائصها والدرجات عليها ، وسيتم الحديث هنا حول النظرية الكلاسيكية

وتعتبر نظرية القياس التقليدية المدخل الرئيسي والوحيد للقياس على مدى قرن من الزمان ، حيث ظهرت ثاني محاولة لظهور ثاني أساس نظري لعملية القياس في بداية الستينات من القرن الماضي ، فقد لعبت الدور الرئيسي والأهم من خلال صياغة المفاهيم الرئيسية لعملية القياس ، إضافة إلى فتح آفاق جديدة للتفكير في منحى نظري آخر يوفر المزيد من الدقة والموضوعية في القياس خاصة القياس النفسي.

والنظرية التقليدية هي واحدة من نظريات القياس التي تستخدم بغرض تحديد العوامل التي تؤثر على الدرجة التي يحصل عليها الفرد في الاختبار وترتكز هذه النظرية على مفهوم الدرجة الحقيقية والدرجة الملاحظة والدرجة الخطأ ، والذي يفترض أنه لو أمكن أن نجرى الاختبار عدة مرات على الفرد بعناصر جديدة وتحت ظروف مختلفة ، فإننا نحصل على درجات ملاحظة مختلفة متوسطها هو أقرب تقدير غير متحيز لقدرة الفرد أو درجته الحقيقية (Randall,1998:4)

وتشكل النظرية الكلاسيكية (التقليدية) أسلوبا بسيطا بين العوامل المؤثرة في الدرجة الظاهرية للمفحوص والتي تسبب ما يسمى بأخطاء القياس ويوضح ذلك من خلال الافتراضات الخاصة به ، إذ يقوم هذا الأسلوب على سبعة (٧) افتراضات ، بمعنى أن دقة وصحة النتائج التي يمكن الحصول عليها تبقى مرهونة بمدى تحقق تلك الافتراضات ، وان أي انتهاك لأي من هذه الافتراضات يعني أن النتائج ستكون موضع الشك. أما الافتراضات التي تقوم عليها هذه النظرية فهي على النحو التالي:

الافتراض الأول:

ويتعلق بمكونات الدرجة الحقيقية للمفحوص والتي تمثل قدرته الحقيقية وينص هذا الافتراض على أن الدرجة الحقيقية لمفحوص ما تتكون من جزأين . الأول يمثل الدرجة الظاهرية أو الملاحظة والثاني الدرجة الخطأ أو خطأ القياس للدرجة الحقيقية ويمكن صياغة هذا الافتراض بالرموز وفقا للمعادلة رقم (١٥) على النحو التالي:

د ح = د ظ + د خ .... (١٥).

ولتوضيح كيفية البساطة التي أظهرتها النظرية الكلاسيكية بالنسبة للدرجة الحقيقية دعنا نفترض أن عددا من المفحوصين خضعوا لاختبار يقيس القدرة العقلية العامة وفي أكثر من مرة حسب ما هو واضح في الجدول رقم (٣)

جدول (٣) مكونات الدرجة الحقيقية حسب النظرية الكلاسيكية

رقم الطالب	التطبيق ١	التطبيق ٢	التطبيق ٣	التطبيق س
م١	د ١،١	د ١،٢	د ١،٣	د س
م٢	د ٢،١	د ٢،٢	د ٢،٣	د ...
م٣	د ٣،١	د ٣،٢	د ٣،٣	د ...
م ن	د ن١	د ن٢	د ن٣	د ن س

ومن هنا فان الدرجة الحقيقة هي درجة نظرية ، والسبب في ذلك أن لكل مرة يخضع فيها المفحوص للاختبار (موقف) أثر على الدرجة مثل: الصعوبة ، نوع الفقرات، عدده ..(اثر الموقف) وكذلك الفرد ذاته غير مستقر ومتغير مثل: القلق الاستعداد، ظروف التطبيق ، المراقب كما أن اثر الاختبار قد يتفاعل مع اثر الفرد نفسه ، كما أن هناك خطا قد يتسبب فيه الاختبار باعتباره أداة قياس وهذا ما نسميه بخطأ القياس ، بمعنى أن هناك أربعة مصادر للأخطاء ولذلك إذا أردنا أن نعرف مكونات الدرجة الظاهرية فإن النموذج العام لهذه الدرجة سيكون على النحو التالي:

د م د = س + ث ت + ث ن + ث ع ن + ث ع حيث :

د م ن . الدرجة الحقيقية للمفحوص س على اختبار يقيس السمة

س : الوسط الحسابي للدرجات الظاهرية التي يحصل عليها في تلك المواقف.

ث ت : اثر الاختبار ونقصد به صورة الاختبار(عدد الفقرات، أنواعها، الدرجة المخصصة لها غرضها.....الخ

ث ن . اثر خصائص الفرد نفسه (القلق، الاستعداد، الدافعية ، الحساسية.. .الخ.

ث ع ن : اثر التفاعل بين الاختبار وخصائص الفرد مثل (الحساسية لنوع الفقرات، تفضيله لنوع معين منها..الخ

ث ع . اثر الخطأ والذي يعود لإجراءات القياس كالتصحيح وتشدد أو تهاون المصحح ......الخ.

وإذا ما أمعنا النظر في هذه المكونات فان هناك شيئا من التعقيد ولذلك جاءت النظرية الكلاسيكية وبسطت هذه المكونات باعتبار أن مجموع كل من أخطاء الاختبار وأخطاء التفاعل وأخطاء الفرد يساوي الصفر وأبقت على كل من الدرجة الحقيقة الخاصة بالمفحوص ودرجة الخطأ.ولذلك أصبح النموذج العام للدرجة الظاهرية لمفحوص ما في سمة ما أصبحت على النحو التالي.

س د = س + ثع حيث :

د س ن : الدرجة الحقيقية للمفحوص س على اختبار يقيس السمة

س . الوسط الحسابي للدرجات الظاهرية التي يحصل عليها في تلك المواقف.

ث خ : اثر الخطأ والذي يعود لإجراءات القياس كالتصحيح وتشدد أو تهاون المصحح . .. الخ

الافتراض الثاني:

إذا خضع المفحوص لاختبار ما يقيس سمة معينة كالتحصيل أو القدرة العقلية فان قدرة الفرد المتعلقة بتلك السمة تتمثل في الدرجة التي سيحصل عليها نتيجة

لإجابته على عدد من الفقرات التي يتكون منها الاختبار أو على جزء منها لكن هذه الدرجة تعتبر درجة ملاحظة أو ظاهرية (Observation Score) أما الدرجة الحقيقية للمفحوص فقد أشارت لها النظرية التقليدية من خلال احد افتراضاتها الذي ينص على " **الدرجة الحقيقية هي المتوسط الحسابي للتوزيع النظري للدرجات الظاهرية التي يمكن أن يحصل عليها المفحوص إذا خضع لاختبار لعدة مرات أو لعدة نماذج اختبارية متوازية (Paralleled) منه وفقراتها مستقلة وتقيس نفس السمة** " وهذا من الناحية العملية أمر صعب المنال ، حيث انه من الصعب ان يحصل المفحوص على نفس الدرجة في كل مرة، إضافة إلى أن الاختبارات المتوازية وحتى المتكافئة (Equivalent) لها شروطها من حيث توزيع الدرجات عليها من حيث الوسط الحسابي وتباين الدرجات وانحرافها المعياري ومعاملات الالتواء والتفلطح. الخ ومن هنا يعتبر تعريف النظرية الكلاسيكية للدرجة الحقيقية تعريف نظريا . وبالرمز فان الدرجة الحقيقة للمفحوص تكتب على الشكل التالي كما هو في المعادلة (١٦).

د ح = س ١+ س ٢+ س٣+ . س ن + هع .. . (١٦).

د ح = س‾ + هع ... .(١٦). حيث:

د ح :الدرجة الحقيقية ، س‾ الوسط الحسابي لدرجات المفحوص في مرات التطبيق ، هع درجة الخطأ

لكن هذا التعريف بالطبع تبين فيما بعد وكما جدته النظرية الحديثة يعاني من بعض الانتقادات حيث أن حصول مفحوص ما على درجة عالية في اختبار سهل لا يعني انه سيحصل نفس الدرجة في اختبار آخر موازي بسبب صعوبة الفقرات والسبب في ذلك أن فقرات الاختبار غير مستقلة عن بعضها البعض.ولتوضيح الدرجة الخطأ دعنا نفترض أن طول تامر الحقيقي (١٦٥ سم) في حين انه عندما تم قياسه من معلم التربية الرياضية كان طوله (١٦٧سم) وعندما تم قياسه من قبل معلم الرياضيات كان (١٦٤سم) فان هناك خطأ في القياس الأول بلغ (٢ سم) وهو خطأ ايجابي بينما هناك خطا في القياس الثاني بلغ (١سم) وهو خطا سلبي لاحظ أن الطول (١٦٥) هو الحقيقي والقياسين (١٦٧ سم ، ١٦٣ سم) هي القياسات الظاهرية ، وكذلك إذا

كانت الدرجة الحقيقية لتحصيل تامر في رياضيات الصف السادس هي (٨٥) وخضع لاختبارين يقيسان نفس المحتوى فحصل على درجة (٧٩) في الاختبار الأول وعلى (٨٧) في الاختبار الثاني ، فان هناك أخطاء تتمثل في الفروق بين الدرجة الحقيقية المفترضة (٨٥) والدرجات في كل من الاختبارين الأول والثاني ومقدارهما (-٦) و(+٢) على التوالي . لاحظ أن الدرجة الظاهرية ستتكون من الدرجة الحقيقية مضافا إليها الدرجة الخطأ ايجابية كانت أو سلبية.ومن المهم الوعي بان الدرجة الحقيقية هي درجة افتراضية حيث لا يمكن أن نخضع المفحوص لعدد لا نهائي من الاختبارات بسبب الوقت الكبير للذلك كذلك الجهد المترتب على ذلك سواء من قبل من يطبق الاختبار (بناء، طباعة، إخراج، بيئة) أو من قبل المفحوص ذاته حيث يصبح التعب عاملا سلبيا لدرجة المفحوص

الافتراض الثالث:

يتعلق الافتراض الثالث بالعلاقة بين الدرج الحقيقية ودرجة الخطأ ، حيث ينص هذا الافتراض على **إن قيمة معامل الارتباط بين الدرجة الحقيقية والدرجة الخطأ تساوي صفر** وبالرموز فان :

ر د ح .د خ = صفر ....(١٧)

وهذا يعني أن أخطاء القياس المنتظمة لدرجات (قدرة)المفحوصين ذوي الدرجات العالية سواء أكان ايجابيا أو سلبيا لن تكون أكبر منها بالنسبة للمفحوصين ذوي الدرجات (القدرة) المتدنية ، فإذا خضع مجموعة من المفحوصين لاختبار ما ومارس المفحوصين الغش وأجابوا على فقرات الاختبار كما يجيب عليها المفحوصين ذوي القدرات العالية أو نتيجة لفهم المفحوصين متدنيين التحصيل لبعض الفقرات نتيجة لقربهم من المعلم حين قراءة وتوضيح الفقرات وعدم فهم المفحوصين من ذوي القدرة العالية نتيجة لبعدهم عن المدرس ولنفس الغرض (نتيجة لعدم سماع المعلم) ، فان طبيعة الدرجات التي سيتم الحصول عليها ستؤدي إلى معامل ارتباط سلبي (عكسي) بين الدرجات الحقيقة ودرجات الخطأ أو معامل ارتباط ايجابي (طردي) في

حال حدث العكس وهذا يعني انتهاك لمضمون هذا الافتراض.وبالطبع يتعلق هذا الافتراض بالاختبار الواحد.

الافتراض الرابع:

ويتعلق هذا الافتراض بدرجات الخطأ لصفة اختبارات حيث يشير هذا الافتراض إلى " أن درجات الخطأ لمفحوص ما خضع لاختبار ما لا ترتبط سلبيا أو ايجابيا بدرجات الخطأ لنفس المفحوص إذا خضع لاختبار آخر" وبالرموز فان

ر دخ١ .دخ٢ = صفر..........................(١٨)

وهذا منطقي إذا ما توفرت متطلبات وظروف التطبيق الجيد وتشابهت في مرات التطبيق ، أما في حال تأثر الدرجات بعوامل مباشرة كالإرهاق في الاختبارات الطويلة أو التهوية والتدفئة أو البنية الشخصية فان مثل هذا الافتراض يعود غير منطقيا أو معقولا، ولذلك ولتحقيق هذا الافتراض وعدم انتهاكه لا بد من الاهتمام بتوفير الظروف والمتطلبات التي يحتاجها التطبيق المناسب للاختبارات لان ذلك سيعمل على تقليل درجات الخطأ في كل تطبيق وهذا يؤدي إلى تقليل معامل الارتباط وكلما زاد الاهتمام بظروف التطبيق كلما آل الارتباط إلى الصفر، حيث يعتبر ذلك متطلب لتحقق هذا الافتراض .

الافتراض الخامس:

يشير هذا الافتراض إلى عدم ارتباط درجة الخطأ في اختبار ما بالدرجة الحقيقة على اختبار آخر يقيس نفس السمة حيث يشير الافتراض إلى أن "الدرجة الخطأ على اختبار ما لا ترتبط بالدرجة الحقيقية على اختبار آخر" وبالرموز فان

ر دخ١ .د ح٢ = صفر......................(١٩)

ومن هنا فان هذا الافتراض قد ينتهك إذا كان احد الاختبارات يقيس سمة شخصية أو بعض مكونات القدرة التي تؤثر في الأخطاء (زيادة أو نقصانا) في الاختبار الآخر ، وإذا ما أمعنا النظر في الافتراضات الخمسة السابقة نجد أنها توصف(تتعلق) درجات الخطأ أو أخطاء القياس ، حيث أشارت إلى أن خطأ القياس غير منتظم، ،

أي أن خطأ القياس هو انحراف عشوائي للدرجة الملاحظة للمفحوص عن الدرجة النظرية المتوقعة للمفحوص، وهذا يعني أن أخطاء القياس لا تشمل الأخطاء المنتظمة.

الافتراض السادس:

يشير هذا الافتراض إلى مواصفات الاختبارات أو النماذج المتوازية لاختبارات تقيس سمة واحدة، ولذلك يسمى أحيانا بافتراض التوازي للاختبارات، حيث إن تصميم نماذج متعددة لاختبارات تقيس سمة واحدة ليس بالضرورة أن تحقق خاصية التوازي ومن هنا فان مضمون هذا الافتراض يبين شروط الاختبارات أو النماذج المتوازية وهي على النحو الذي سنبينه بعد قليل ، فإذا كان ت١، ت٢) اختبارين يقيسان سمة واحدة كان تكون القدرة العقلية ، وحتى نعتبر هذين الاختبارين متوازيين لا بد من توافر الشروط التالية.

- تشابه المحتوى الذي يقيسه كل من النموذجين.
- تطابق توزيع الدرجات الظاهرية(الملاحظة)لكلا النموذجين وهذا يعني.
  - ✓ تساوي الدرجات في كلا النموذجين. أ ي أن $ت_1 = ت_2$
  - ✓ تساوي الوسط الحسابي للدرجات الظاهرية لكل من النموذجين. أي أن $\bar{س}_1 = \bar{س}_2$.
  - ✓ تساوي تباين الدرجات الظاهرية لكل من النموذجين أي أن $ع^2_1 = ع^2_2$.
  - ✓ تساوي معاملات الالتواء والتفلطح لكلا النموذجين أي أن $م_{ر1} = م_{ر2}$ وكذلك $م_{ط1} = م_{ط2}$.
  - ✓ الرتب المئينية للدرجات على كل من النموذجين متساوية أي أن $م_{ن1} = م_{ن2}$

وإذا ما أمعنا النظر في مضمون الافتراضات السابقة والوقوف على طبيعة الواقع العملي لبناء وتطبيق الاختبارات والمقاييس التي تقيس السمات المختلفة معرفية أو وجدانية أو نفسحركية نجد انه من الصعب من تحقق الافتراضات بشكل متكامل كما هي بشكلها النظري وذلك لعدة عوامل قد تتعلق بأداة القياس أو بطبيعة السمة

المقاسة أو بالمفحوص أو بالتطبيق والتصحيح أو حتى بالأسلوب الإحصائي المستخدم في تحليل النتائج، ولذلك فان أي عملية قياس وخاصة لتلك السمات التي تقع على مستوى قياس فئوي أو أدني يتنابها الشك بمقدار أخطاء القياس التي تقع نتيجة لعملية القياس والتي تشكل الفرق بين الدرجة الحقيقة والدرجة الملاحظة، وبسبب تشدد شروط الاختبارات المتوازية تم اللجوء إلى الاختبارات المتكافئة التي سنتحدث عنها في الافتراض التالي. .

الافتراض السابع:

ويتعلق هذا الافتراض بنماذج الاختبارات المتكافئة (Equivalent)من حيث شروط التكافؤ ، وحيث أن اعتبار تحقق خاصية التكافؤ قد لا تعكس الواقع الحقيقي لقدرة مفحوص ما فقد تم تحديد مواصفات أو شروط لتلك الاختبارات ولاعتبار أن تكافؤ اختبارين أو أكثر لا بد من توفر الشروط التالية

- تشابه المحتوى الذي يقيسه كل من النموذجين.
- بناء الاختبارات بنفس المواصفات (العدد، الإجراءات، النوع ...)
- ليس بالضرورة أن تساوى الدرجات الحقيقية في كلا النماذج.
- ليس بالضرورة تساوي الوسط الحسابي للدرجات الظاهرية لكل من النموذجين. أي أن $\bar{س}_1 = \bar{س}_2$.
- تساوي تباين الدرجات الظاهرية لكل من النموذجين. أي ان $ع^2_1 = ع^2_2$.
- التباينات المشتركة لهذه الاختبارات لا بد أن تكون متساوية. بمعنى أن.

$ع^2_{1\cdot2} = ع^2_{1\cdot3} = ع^2_{2\cdot3}$ وهكذا حسب عدد الاختبارات أو النماذج.

ويمكن تلخيص جوهر النظرية التقليدية من خلال افتراضاتها ، حيث تتكون الدرجة الظاهرية (الملاحظة) من جزأين :الأول الدرجة الحقيقية (ت) والثاني : خطأ عشوائي يسمى درجة الخطأ، حيث لا يرتبط الجزأين معا أو أن قيمة معامل الارتباط بينهما يساوي صفر سواء لاختبار واحد ، وكذلك الحال بينهما في حال تطبيق أكثر من اختبار .كما تشير النظرية إلى مفهوم النماذج المتوازية من الاختبارات

التي تقيس سمة معينة وتحدد مجموعة من الشروط لتحقق افتراض التوازي كما تشير النظرية مفهوم الاختبارات المتكافئة والتي تتحرر من بعض شروط الاختبارات المتوازية نتيجة لصعوبة توفر تلك الشروط في الواقع العملي.كما تعتبر كل من الدرجة الحقيقية والأخطاء الخاصة بها درجات نظرية لا من الصعب الوصول إليها، وإنما يتم تقديرها من الدرجة الملاحظة ، وذلك من خلال حساب المتوسط الحسابي للدرجات الملاحظة التي يحصل عليها المفحوص نتيجة لخضوعه لاختبار يقيس السمة المارد قياسها لعدد لا نهائي من المرات وتلذلك فان الدرجة الملاحظة للمفحوص على اختبار أو عدة اختبارات تقيس سمة ما ليس من الضرورة أن تعكس الدرجة الحقيقية للمفحوص في تلك السمة .

الاستنتاجات

من خلال استعراض افتراضات النظرية التقليدية بمضمونها النظري والمعادلات المصاغة الخاصة بكل منها فانه يمكن الخروج بالعديد من الاستنتاجات الفرعية التي توجه إجراءات النظرية الخاصة ببناء أدوات القياس وتطبيقها وتحليل الدرجات عليها والتنبؤ بالدرجات الحقيقة للمفحوص في مختلف السمات ، أما هذه الاستنتاجات فهي على النحو التالي:

- مجموع درجات الخطأ يساوي صفر ، حيث يكون بعض هذه الأخطاء موجب (أعلى من الدرجة الحقيقة) وأحيانا تكون سلبية (اقل من الدرجة الحقيقية) وعند جمع هذه الأخطاء فإنها ستؤول إلى الصفر
- مجموع حواصل ضرب الدرجة الخطأ في الدرجة الحقيقية يساوي صفر أي أن مجموع (د.خ .د.ح) = صفر، وكذلك فان التباين المشترك بينهما يساوي صفر حيث انه عبارة عن حاصل ضرب مجموع الدرجات الخطأ في مجموع الدرجات الحقيقية أي أن

$$ع_{ح خ} = \text{مجموع (د.خ .د.ح)} - \text{مجموع د.خ} \times \text{مجموع د.ح}$$

- تباين الدرجات الظاهرية يكون مساويا لتباين الدرجات الحقيقية وتباين الدرجات الخطأ، أي أن :

$$ع^{2}_{د.م} = ع^{2}_{د.ح} + ع^{2}_{د.خ} \quad ....(٢)$$

مع تـذكر أن تبـاين الدرجـة الخطـأ بالنسـبة للاختبـارات المتوازيـة يكـون متسـاويا.بمعنى انـه إذا تـوفرت قـيم تبـاين الـدرجات الخطـأ والـدرجات الظاهريـة والدرجات الحقيقية لكل المفحوصين فان تبـاين الدرجـة الظاهريـة لا بـد أن يسـاوي مجموع تبايني كـل مـن الدرجـة الحقيقـة والدرجـة الخطأ.وقـد يكـون تبـاين الدرجـة الظاهرية مـساويا لتباين الدرجة الحقيقية ، وذلك في حال عدم وجود أخطاء قياس أي أنه إذا كان $\text{ع}^{2}_{\text{خ}}$ = صفر فان $\text{ع}^{2}_{\text{م}} = \text{ع}^{2}_{\text{ح}}$ أي أن الفروق بين الدرجة الظاهرية تساوي الفروق في الدرجة الحقيقية وإذا كان كانـت عمليـة القيـاس تتابهـا الأخطـاء (عدم الدقة) فان تباين الدرجة الخطأ يساوي قيمة معينة أكبر من صـفر وكـذلك فـان الفروق في الدرجة الظاهرية للمفحوصـين سـتعكس الفـروق بـين الـدرجات الحقيقـة والدرجات الخطأ، وإذا كان تباين الدرجات الخطأ كبيرا فان ذلك سينعكس على كـل من تباين الدرجات الظاهرية وتباين الدرجات الحقيقية حيث انه إذا كانت الفروق بين الدرجات الظاهرية سيكون سببه أخطـاء القيـاس أكثـر منـه فـروق نتيجـة لفـروق في الدرجات الحقيقية، أي أن أخطاء القياس تزداد بزيادة التباين في الدرجات الظاهرية.

- معامل الارتباط بين الدرجات الحقيقيـة والظاهريـة يسـاوي الانحـراف المعيـاري للدرجات الحقيقية مقسـوما على الانحراف المعيـاري للـدرجات الظاهريـة وهـذا يعني أن النسبة بين تباين الـدرجة الحقيقية إلى تباين الدرجة الظاهرية يساوي مربع معامل الارتباط بين كل من الدرجة الظاهرية والدرجة الحقيقية أي أن

$$\text{ر}^{2}_{\text{م.ح}} = \text{ع}^{2}_{\text{ح}} / \text{ع}^{2}_{\text{م}} \quad (21)$$

وهذا يعني أنه يمكن وصف نسبة التباين في الدرجات الحقيقية من خلال علاقة خطية مع الدرجات الظاهرية ، وانه يمكن أن تكون الـدرجات الظاهريـة متنبئ جيد بالـدرجات الحقيقية عندما تكون قيمة الارتباط الخطي تامة أي عنـدما تكـون قيمة $\text{ر}_{\text{م.ح}} = 1$ أي عندما يكون الانحراف المعياري لكـل مـن الـدرجات الظاهريـة والحقيقيـة متسـاويين. وكذلك فان قيمة مربع معامل الارتباط تساوي قيمة معامـل الارتبـاط التـام مطروحـا منها نسـبة التباين لكل من الدرجات الخطأ إلى تباين الدرجات الظاهرية، وبالرموز فان

$$\text{ر}^{2}_{\text{م.ح}} = 1 - \text{ع}^{2}_{\text{خ}} / \text{ع}^{2}_{\text{م}} \quad (22)$$

وهذا يعطينا أكثر من بديل لتفسير قيمة معامل الارتباط بين الدرجات الظاهرية والدرجات الحقيقية، حيث انه إذا كانت قيمة تباين الدرجات الخطأ منخفض ومقاربة لقيمة تباين الدرجات الظاهرية فان نسبة التباين المفسر في الدرجات الحقيقة من خلال الدرجات الظاهرية يكون عاليا وحينها سيكون الدرجات الظاهرية متنبئ قوي بالدرجات الحقيقية ، كذلك إذا كانت قيمة تباين الدرجات الخطأ عالية ومقاربة لقيمة تباين الدرجات الظاهرية فان نسبة التباين في الدرجات الحقيقة من خلال الدرجات الظاهرية يكون متدنيا وعندها سنكون الدرجات الظاهرية متنبئ غير جيد للدرجات الحقيقية

- تباين الدرجات الظاهرية على نموذجين متوازيين لاختبار ما يقيس سمة معينة لا بد أن يكونا متساويين، وإلا فان شرط مهما من شروط التوازي قد انتهك أو لم يتحقق.
- تتساوى قيم معاملات الارتباط بين الدرجات على نماذج متوازية من اختبار يقيس سمة معينة ، أي انه إذا كانت س١ درجة على نموذج ما وس٢ درجة على نموذج آخر موازي للنموذج الأول فان قيمة معامل الارتباط بين كل من س١ أو س٢ مع درجة أخرى (ص) تكون متساوية ، أي أن ارتباط الدرجات على النماذج المتوازية مع أي درجة أخرى تكون متساوية .بمعنى أن:

$$ر_{\text{س١.ص}} = ر_{\text{س٢.ص}} \qquad (٢٣)$$

- قيمة معامل الارتباط بين الدرجات الظاهرية على نموذجين متوازيين تكون مساوية لنسب التباين الحقيقي إلى التباين الظاهري للدرجات على كلا النموذجين.أي أن :

$$ر_{\text{م١م٢}} = ع^{٢}_{\text{ح١}} / ع^{٢}_{\text{م١}} = ع^{٢}_{\text{ح٢}} / ع^{٢}_{\text{م٢}} \qquad (٢٤)$$

وهذا يعني أن الارتباط بين الدرجات على كلا الاختبارين ارتباط تام اي أن قيمته تساوي (١) وهذا لا يتحقق إلا إذا كانت النسب أعلاه تساوي (١) ، وهذا

بالطبع لا يحدث إلا إذا كانت عملية القياس دقيقة بحيث لا يتخللها أخطاء أي عندما تكون قيمت تباين الدرجات الخطأ تساوي صفر وعندها يكون تباين الدرجات الحقيقية مساويا لتباين الدرجات الظاهرية على بالنسبة للنموذجين.

- قيمة معامل الارتباط بين الدرجات على نموذجين متوازين لاختبار ما يقيس سمة واحدة يساوي قيمة الارتباط التام مطروحا منه نسبة تباين الخطأ إلى تباين الدرجات الظاهرية اي أن:

$$ر_{م١.م٢} = ١ - ع^٢_{دخ} / ع^٢_{دم} \quad \ldots\ldots \quad (٢٥)$$

وهذا يعني أن الارتباط بين النماذج المتوازية ارتباط تام أي ر = ١ ، ويتحقق ذلك عندما تكون قيمة تباين الخطأ مساوية للصفر وكذلك يكون الارتباط مساوياً للصفر عندما يتساوى كل من تباين الخطأ والتباين الظاهري متساويين

- قيمة معامل الارتباط بين الدرجات على نموذجين متوازين من الاختبارات لنفس السمة يساوي قيمة الارتباط التام مطروحا منه مربع قيمة معامل الارتباط بين الدرجات الظاهرية والدرجات الخطأ وبالرموز فان :

$$ر_{م١.م٢} = ١ - ر^٢_{(دم\ دخ)} \quad \ldots\ldots \quad (٢٦)$$

حيث(م١. م٢) درجات على اختبارين متوازين وكذلك (دم دخ) الدرجات الملاحظة ودرجات الخطأ على التوالي

وتكون هذه القيمة مساوية للقيمة (١) عندما لا يكون ارتباط بين الدرجات الظاهرية والدرجات الخطأ ، ويكون الارتباط بين النماذج المتوازية مساويا للصفر عندما ترتبط الدرجات الملاحظة أو الظاهرية بدرجات الخطأ ارتباطا قويا أو عاليا.

- مربع معامل الارتباط بين الدرجات الظاهرية والدرجات الحقيقية يساوي الارتباط بين الدرجات الظاهرية على الاختبارات المتوازية، ولذلك إذا تحققت صفة التوازي وبالتالي الحصول على قيمة معامل الارتباط بين الدرجات الظاهرية والحقيقية فانه يمكن تقدير قيمة مربع معامل الارتباط بين الدرجات

الظاهرية والحقيقية على النماذج المتوازية ، وهذه الحقيقية مهمة بسبب عدم معرفة الدرجات الحقيقية وبالتالي عدم القدرة على حساب الارتباط بين الدرجة الظاهرية والدرجة الحقيقية بشكل مباشر .أي أن .

ر٢مح.مح = ر م١.م٢........ ... ... (٢٧)

- التباين الخاص بالدرجة الحقيقية يساوي التباين المشترك بين الدرجات الظاهرية على النماذج المتساوية.ولذلك فان التباين المشترك للدرجات الظاهرية على النماذج المتساوية يمكن أن يكون متنبئا بتباين الدرجة الحقيقية وبالرموز فان ع٢د.ح = ع م .م - حيث ع٢د ح تباين الدرجة الحقيقية وع م .م - التباين المشترك بين الدرجات على نموذجين متوازيين
- تباين الدرجة الخطأ يكون مساويا لتباين الدرجة الظاهرية مضروبا بالقيمة (١- رم م -) حيث رم .م - هو معامل الارتباط بين الدرجات الظاهرية على نموذجين متوازيين. وبالرموز فان:

ع٢مخ = ع٢ دم (١-م م-). . .... (٢٨)

وإذا أخذنا قيمة الجذر التربيعي لتباين الخطأ (ع مخ) فسوف نحصل على ما يسمى بالخطأ المعياري للقياس حيث تستخدم المعادلة الناتجة يمكن أن تستخدم في تقدير كل من التباين والتباين المشترك لدرجات الخطأ حيث :

ع٢مخ = ع٢مخ = ع٢مخ = ع٢دم (١- رم١.م٢). (٢٩).

- الارتباط بين الدرجات الحقيقية على نموذجين متوازيين يساوي الارتباط بين الدرجات الملاحظة (الظاهرية) على النموذجين مقسوما على الجذر التربيعي لحاصل ضرب الارتباط بين الدرجات الظاهرية على نموذجين متوازيين في معامل الارتباط بين الدرجات الظاهرية على نموذجين متوازين آخرين بمعنى انه سيكون في هذه الحالة أربع صور اختبارية ولكن كل نموذجين يوازي منهما الآخر ، وبالرموز فان

$$\text{ر}_{\text{ح}1\text{،ح}2} = \frac{\text{ر}_{\text{م}1\text{،م}2}}{\sqrt{\text{ر}_{(\text{م}1\text{،م}1\text{ه})} \times \text{ر}_{(\text{م}2\text{،م}2\text{ه})}}} \quad \ldots\ldots\ldots\ldots(٣٠)$$

حيث .

$\text{ر}_{\text{ح}1\text{،ح}2}$: معامل الارتباط بين الدرجات الحقيقية على نموذجين متوازيين

$\text{ر}_{\text{م}1\text{،م}2}$: معامل الارتباط بين الدرجات الظاهرية على نموذجين متوازيين.

$\text{ر}_{(\text{م}1\text{،م}1\text{ه})}$: معامل الارتباط بين الدرجات الظاهرية على نموذجين متوازيين

$\text{ر}_{(\text{م}2\text{،م}2\text{ه})}$ معامل الارتباط بين الدرجات الظاهرية على نموذجين متوازيين.

م١ وم١ه هي درجات ملاحظة على اختبارين متوازيين وبالتالي فان.

م١= ح م١+ خ م١ وكذلك فان م١ه = ح.م١ه+ خ م١ه. وكذلك فان:

م٢=ح.م٢+خ.م٢ وكذلك فان م٢ه = ح.م٢ه+ خ.م٢ه.

ولا بد من الانتباه أن القيمة تحت الجذر لا بد وان تكون أكبر أو تساوي القيمة(١) وكذلك القيمة المراد إيجادها من المعادلة (٣) لا بد وان تكون اكبر من قيمة البسط في المعادلة أعلاه حيث يعمل الارتباط بين الدرجات الملاحظة على تخفيض الارتباط بين الدرجات الحقيقية وهوما يسمى بتصحيح قيمة معامل الارتباط لأثر التخفيض ، كما يساعد ذلك في تقدير الارتباطات بين الدرجات الحقيقية على النماذج المتوازية للاختبارات.

- إذا تم تطبيق عدة نماذج من الاختبارات المتوازية لعدة مرات (ن) وكانت س تشير إلى عدد هذه النماذج فان س = مجموع الدرجات على اختبار موازي آخر أي أن

$$\text{س} = \sum_{\text{ف}=1}^{\text{ن}} \text{ص ف}$$

حيث.نموذج اختبار، ن عدد النماذج أو عدد مرات التطبيق ، ص درجات الاختبار

ولذلك إذا كان ص هو نموذج اختباري فان تباين الدرجات الحقيقية ($ع^{2}$ حج س) على النموذج س يساوي تباين الدرجات الحقيقية على النموذج ص($ع^{2}$ حج ص) وإذا تم تقسيم اختبار ما إلى عدة نماذج جزئية فان تباين الدرجات الحقيقية للاختبار ككل تساوي ع٢ حج س عدد مرات النماذج مرفوعا للقوة (٢) مضروبا في تباين الدرجات الحقيقية على تلك النماذج أي ان

$$ع^{2}_{حج س} = ن^{2} ع^{2}_{حج ص} \quad (٣١)$$

وهذه النتيجة كانت مهمة في تطوير معادلة سبيرمان براون الخاصة بمعامل الثبات التي تشير إلى دقة الاختبار.

- إذا كان ص هو نموذج اختباري فان تباين الدرجات الخطأ (ع٢د.خ س) على النموذج س يساوي تباين الدرجات الخطأ على النموذج ص(ع٢د.خ ص) وإذا تم تقسيم اختبار ما إلى عدة نماذج جزئية فان تباين الدرجات الخطأ للاختبار ككل (ع٢ د.خ س) تساوي عدد مرات النماذج مضروبا في تباين الدرجات الحقيقية على تلك النماذج أي أن

$$ع^{2}_{حج س} = ن ع^{2}_{حج ص}$$

ومن خلال النتيجة الحالية والسابقة يتبين أن تقسيم الاختبار إلى عدة أجزاء (نماذج) يساهم في زيادة تباين الدرجات الحقيقية ، وهذا يشير إلى أن زيادة طول الاختبار يساهم في توفير دقة قياس أكثر، بمعنى أن دقة القياس في اختبار ما أكثر من دقة اختبارات جزئية لنفس الاختبار، والسبب يعود إلى أن أخطاء القياس يتم تكرارها بعدد النماذج ويمكن استيضاح ذلك من خلال مراجعة معادلة سبيرمان براون لمعامل الثبات والتي تعبر عن معامل الثبات الذي يشير إلى دقة الاختبار.

عدد مرات مجموع الدرجات الكلي لعدد من المفحوصين على مجموعة من النماذج الاختبارية المتوازية يساوي

- المتوسط الحسابي للدرجات الملاحظة على اختبار ما يقيس سمة معينة يمثل المتوسط الحسابي للدرجات الحقيقة للمفحوص في تلك السمة وبالرموز فان د م = د ح حيث د م الدرجة الملاحظة ، دح الدرجة الحقيقية
- قيمة معامل الارتباط بين الدرجة الحقيقية ودرجة الخطأ تساوي صفر ، بمعنى انه لا يوجد ارتباط بينهما ر دح .د خ = صفر وهذا يعني بان تحديد الدرجة الحقيقية لا يعني بالضرورة تحديد الدرجة الخطأ ا وان تحديد الدرجة الخطأ لا يعني بالضرورة تحديد الدرجة الحقيقية أيضا ، أي أن ر دح .د خ = صفر وكذلك ر دح١ .د خ ٢= صفر.
- قيمة معامل الارتباط بين الدرجات الظاهرية على اختبارين متوازيين تساوي حاصل قسمة قيمة تباين الدرجات الحقيقية على قيمة تباين الدرجات الظاهرية. أي أن ر دم .د = ع٢ د.م / ع٢ د.ح .
- إن التباين المشترك لتباين نموذجين اختباريين متوازيين تساوي التباين الحقيقي أي أن : ع٢١ = ع٢٢ .
- إن القدرة التنبوية لمجموعة من نماذج الاختبارات المتوازية متساوية بمعنى أنه يمكن تحديد تباين الخطأ من خلال عدة طرق

الفصل الثاني

مفهوم الثبات

مقدمة

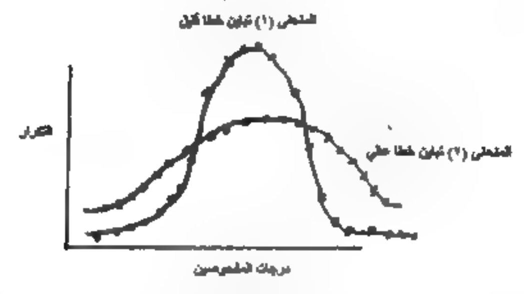
يحتاج مستخدم أو مطور الاختبار عادة إلى الثقة بنتائج الاختبار حيث تعتمد عليه الكثير من القرارات ، حيث يشير مفهوم الثبات إلى دقة الاختبار وبالتالي دقة القرارات المتخذة اعتمادا على نتائجه ، ويشار إلى دقة الاختبار بمؤشر يسمى مؤشر الثبات أو معامل الثبات ، ومن الجدير بالذكر أن الدقة ليست للاختبار بقدر ما هي لنتائج الاختبار فالأصح أن نقول ثبات نتائج الاختبار أو درجات الاختبار ، ويعرف الثبات في الواقع من خلال العديد من الصور والتي سبق وتم الإشارة لها بطريقة غير مباشرة من خلال الاستنتاجات الخاصة بافتراضات النظرية التقليدية في القياس **(انظر الفصل الأول)** ويمكن اعتماد احد ستة تعريفات وهي على النحو التالي:

- معامل الثبات (م ث) لاختبار ما هو معامل الارتباط (ر م.م*) بين الدرجات الظاهرية على هذا الاختبار والدرجات الظاهرية على اختبار موازي آخر يقيس نفس السمة، وهذا يعني انه إذا كانت قيمة معامل الارتباط بين درجات المفحوصين على النموذجين المتوازيين عالية (ر م.م*)= ١ فمعنى ذلك أن الأختبار يتميز بالثبات العالي (م ث) = ١ .وإذا كانت قيمة معامل الارتباط منخفضة فان الاختبار يكون قليل الثبات ، وفي الحالتين يكون الدليل على دقة الاختبار كأداة قياس. أي أن (م ث)= (ر م.م*).
- قيمة مربع معامل الثبات (م٢ث) تعبر عن نسبة التباين المفسر من خلال العلاقة الخطية بين الدرجات الملاحظة على اختبارين متوازيين.
- معامل الثبات هو مربع معامل ارتباط بيرسون (ر م.م*) بين الدرجات الظاهرية على نموذجين متوازيين يقيسان نفس السمة، وهنا يشير معامل الثبات إلى نسبة

التباين المفسر لقدرة الطلبة من خلال الارتباط بين درجاتهم على النموذجين ومن هنا فان معامل الثبات يساوي مربع معامل الارتباط هو(ر٢م م/) حيث م، م/ هما الدرجات الملاحظة على اختبارين متوازيين

* معامل الثبات هو نسبة تباين الدرجات الحقيقية إلى تباين الدرجات الملاحظة ، وبما أن معامل الثبات هو قيمة تتمثل في عدد حقيقي بين (صفر-١)، ويكون الثبات عاليا إذا كانت قيمته = (١) أي أن قيمة (ع٢ح. / ع٢م) = ١ ولذلك فان أي تباين في الدرجات الملاحظة للمفحوصين تعكس بالضرورة تباينا في الدرجات الحقيقية أكثر منه تباينا في الأخطاء كذلك إذا كانت قيمة معامل الثبات (م ث) = ١ أي القيمة (ع٢ح / ع٢م) تساوي (١) فان أية فروق في الدرجات الملاحظة للمفحوصين لا بد أن تعكس فروقا في درجاتهم الحقيقية وإذا تساوى كل من تباين الدرجات الملاحظة وتباين الدرجات الحقيقية اي أن ع٢ ح =ع٢ م فان تباين الخطأ في هذه الحالة لا بد وان تساوي صفر أي انه لا توجد أخطاء قياس ، ومن هنا فان معامل الثبات هو مؤشر على دقة القياس

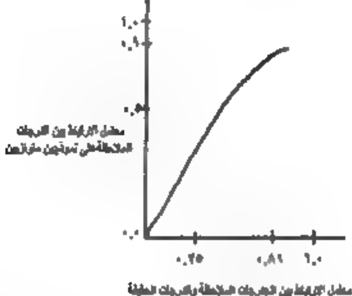
أما إذا كانت قيمة الثبات اقل من (١) فان ذلك يعني أن هناك أخطاء قياس، وعندما يكون الثبات صفر فان ذلك يعني أن تباين الدرجات الملاحظة سيساوي تباين الدرجات الخطأ وهذا يعني كل الدرجات تتضمن أخطاء قياس ولذلك فان الفروق بين الدرجات الملاحظة للمفحوصين تعكس أخطاءً عشوائية أكثر منها فروقا في الدرجات الحقيقية. ولذلك فانه كلما زادت قيمة معامل الثبات فان قيمة تباين الخطأ تقل، وعندما تكون قيمة تباين الأخطاء قليلة بحيث يمكن تجاهلها فان الدرجات الملاحظة للمفحوصين تكاد تكون مساوية أو قريبة من الدرجات الحقيقية ، وعلى العكس فانه عندما تزيد قيم تباين الخطأ فان الدرجات الملاحظة تكون ضعيفة التقدير للدرجات الحقيقة ، ويمكن استيضاح ذلك من خلال الشكل (٧) ، حيث تمثل المنحنيات التوزيع النظري للدرجات الملاحظة لدرجة حقيقية محددة لأحد المفحوصين.



شكل (٧) العلاقة بين الارتباط بين الدرجات الملاحظة والحقيقية على الاختبار وأثرها على ثبات الاختبار

من خلال الشكل يظهر الدرجات الملاحظة تقترب من الدرجات الحقيقية ويحدث ذلك عندما يتساوى كل من تباين الدرجات الملاحظة وتباين الخطأ وفي نفس الوقت عدم وجود تباين في الدرجات الحقيقية أي ع٢ح =صفر، ع٢م = ع٢ع .كما في المنحنى (١) حيث تقترب الدرجات الحقيقية من الدرجات الملاحظة، أما عندما يكون تباين الخطأ كبيرا كفي المنحنى (٢) فان الدرجات الملاحظة تكون بعيدة عن الدرجات الحقيقية وهذا مؤشر على انخفاض ثبات الاختبار

- معامل الثبات لاختبار ما يساوي قيمة مربع معامل الارتباط بين الدرجات الملاحظة والدرجات الحقيقية ، فإذا كان معمل الثبات م ث = (٠.٨١) فان معمل الارتباط بين الدرجات الملاحظة والدرجات الحقيقية = (٠.٩) وتتضح هذه العلاقة من خلال الشكل (٨) أدناه حيث يتبين انه إذا كان معامل الثبات اكبر من الصفر واقل من الواحد الصحيح (١ < م ث > صفر) فان معامل الارتباط بين الدرجات الملاحظة والحقيقية (ر م.خ) سيكون أكبر من معامل الثبات مث وهنا سترتبط الدرجات الملاحظة بالدرجات الحقيقية الخاصة بها بشكل اكبر من ارتباطها بالدرجات الملاحظة على اختبار متوازي آخر.

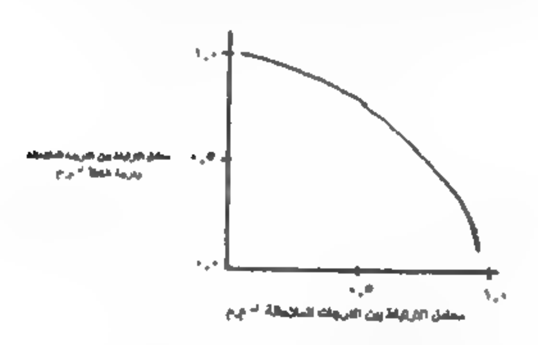


شكل (٨) العلاقة بين الارتباط الدرجات الملاحظة والحقيقية والدرجات الملاحظة والحقيقية

وبما أن ارتباط درجة الاختبار مع الدرجة الحقيقية أعلى ارتباطها مع أي متغير آخر ، فان الحد الأعلى لقيمة معامل ارتباط الدرجات الملاحظة مع أي متغير آخر لن تتعدى قيمة الجذر التربيعي لمعامل الثبات أي أن وإذا ما أردنا استخدام الدرجات الملاحظة (د م) في التنبؤ بالدرجات الملاحظة (د م•) على اختبار آخر وتم حساب قيمة معامل الارتباط بين الدرجات الملاحظة على كلا الاختبارين فان معامل الارتباط (ر م م•) في مثل هذه الحالة يسمى بمعامل الصدق ، وبما أن ارتباط الدرجات الملاحظة بدرجات المحك أي (ر م م•) لن يكون أعلى من ارتباط الدرجة الملاحظة بالحقيقية (ر م خ) فان قيمة (ر م م•) لن تكون أعلى من قيمة الجذر التربيعي لمعامل الثبات ، بمعنى أن الجذر التربيعي لقيمة معامل الثبات هي السقف الأعلى لمعامل الارتباط بين الدرجات الظاهرية للاختبار والدرجات على متغير آخر، ولذلك فان عدم الثبات للاختبار يؤثر في صدق ذلك الاختبار.وعلى الرغم من أن قيمة معامل الصدق لا تتجاوز قيمة الجذر التربيعي لمعامل الثبات ، إلا أن قيمة معامل الصدق يمكن أن تكون أعلى من قيمة معامل الثبات نفسه .وعلى سبيل المثال إذا كانت قيمة

^{2}ن = (٤٩) فستكون قيمة (^{2}م ج) = (..٧) وهنا فان قيمة (^{2} م م) على اعتبار أنها تمثل قيمة معامل الصدق ستكون أعلى من قيمة معامل الثبات نفسه(٤٩.)، وهي هنا (..٧).

- قيمة معامل الثبات للاختبار (م ث) = ١- ر٢م خ أي أن معامل الثبات هو القيمة (١) مطروحا منها قيمة مربع معامل الارتباط بين الدرجات الملاحظة والخطأ، وتتضح العلاقة بين معامل الثبات (م ث) ومعامل الارتباط بين الدرجات الملاحظة ودرجات الخطأ (م م.ح) من خلال الشكل رقم (٩) .حيث يكون الثبات تاما أي (م ث = ١) فقط عندما تكون قيمة (رم خ = صفر) .



شكل(٩) العلاقة بين معامل الارتباط بين الدرجات الملاحظة ومعامل الارتباط بين الدرجة الملاحظة والدرجة الخطأ

- يرتبط الثبات بتباين الدرجات الخطأ وتباين الرجات الملاحظة والعلاقة بينهما ، حيث أن معامل الثبات (م ث = ١-ع٢خ / ع٢ م) فعندما يكون الثبات تاما أي (م ث= ١) فان تباين الخطأ ع٢خ = صفر وعندما ينعدم ثبات الاختبار أي

أن م ث =صفر) فان تباين الأخطاء سيتساوى مع تباين الدرجات الملاحظة ، ويتأثر الاتساق الداخلي بالنسبة للاختبار من مجموعة إلى أخرى حيث قد يقل أو يرتفع الثبات من مجموعة إلى أخرى، وهذا ما يسمى بظاهرة ضيق المدى للدرجات.

وباختصار فانه إذا كانت قيمة معامل الثبات تساوي (١) أي أن ر س = ١ فان هناك مجموعة من الاستنتاجات على النحو التالي

- دقة القياس عالية لدرجة انه لا يوجد أخطاء قياس وان مجموع اخطاء القياس يساوي صفر. (مجموع د.م = صفر.
- الدرجات الملاحظة تساوي الدرجة الحقيقية أي أن (د م = د.ح) لجميع المفحوصين.
- الفروق بين الدرجات الملاحظة تعكس فروق في الدرجات الحقيقية .
- تباين الدرجات الملاحظة يساوي تباين الدرجات الحقيقية أي ع٢ د.م = ع٢ د.ح
- الارتباط بين الدرجات الملاحظة الحقيقية ارتباطا تام أي أن ر م ح = ١.
- لا يوجد ارتباط بين الدرجات الملاحظة والدرجات الخطأ اي أن رم خ = صفر.

وكذلك إذا كانت قيمة معامل الثبات ر س = صفر فان هناك مجموعة من الاستنتاجات على النحو التالي:

- الخطأ الوارد في عملية القياس هو خطأ عشوائي فقط
- الدرجة الكلية تساوي الدرجة الخطأ لكل المفحوصين.
- التباين الدرجات الملاحظة يعكس تباين في درجات الخطأ أي ع٢ د م = ع٢ د.خ
- الفروق بين الدرجات تعكس أخطاء القياس.
- لا يوجد ارتباط بين الدرجات الملاحظة والحقيقية اي أن ر .م.خ = صفر.
- الارتباط بين الدرجات الملاحظة ودرجات الخطأ ارتباطا تام أي أن رم.خ = ١.

دار جرير للنشر والتوزيع

وكذلك إذا كانت قيمة معامل الثبات أكبر أو يساوي (صفر) أو أقل من أو يساوي (١) أي أن (صفر ≤ م ث ≤ ١) فان هناك مجموعة من الاستنتاجات على النحو التالي

- سيتتاب عملية القياس بعض الأخطاء
- الدرجة الملاحظة هي مجموع الدرجة الحقيقية والدرجة الخطأ
- تباين الدرجات الملاحظة سيتضمن جزءا من تباين الدرجات الحقيقية وتباين الدرجات الخطأ أي أن ع٢ م = ع٢ ح + ع٢ خ
- الفروق بين الدرجات الملاحظة تعكس بالضرورة فروقا في الدرجات الحقيقية والدرجات الخطأ
- معامل الارتباط بين الدرجات الملاحظة والدرجات الحقيقية يساوي قيمة الجذر التربيعي لمعامل الثبات للاختبار. رم ح = $\sqrt{\text{م ت}}$
- معامل الارتباط بين الدرجات الملاحظة والدرجات الخطأ يساوي قيمة الجذر التربيعي للقيمة
- (١) مطروحا منها معامل الثبات أي ان رم ح = $\sqrt{\text{١ − م ت}}$
- كمعامل الثبات هو نسبة تباين الدرجات الحقيقية إلى تباين الدرجات الملاحظة أي أن م ث = ع٢ ح / ع٢م.
- كلما زادت قيمة معامل الثبات ازدادت الثقة بإمكانية تقدير الدرجة الحقيقية من خلال الدرجة الملاحظة، وسبب ذلك انخفاض قيمة خطأ القياس.

تقدير معامل الثبات. Reliability

يعد الثبات من العوامل الهامة أو الخصائص الواجب توافرها لصلاحية استخدام أي اختبار أو جهاز قياس ، فالمقياس أو الجهاز الثابت سوف يعطي نفس النتيجة تقريبا لنفس الشخص عند إجراء القياس لمرات عديدة في نفس اليوم أو أيام مختلفة حيث تكون تلك النتيجة مؤشرا جيدا لقدرات هذا الشخص.

ويتعلق الثبات بدقة القياس بصرف النظر عما يقاس وتتضمن جميع القياسات العملية بعض الخطأ العشوائي الذي يؤدي لعدم ثبات النتائج ومن المعاني التي يمكنها مفهوم الثبات.

- إن الاختبار موثوق به ويعتمد عليه ، كما يعني الاستقرار أي انه لو أعيد تطبيق الاختبار نفسه على الفرد الواحد فانه يعطي شيئا من الاستقرار في النتائج
- وهو اتساق الدرجات التي يحصل عليها نفس الأفراد في مرات الأجراء المختلفة
- وهو الاختبار الذي لو أعيد تطبيقه على نفس الأفراد فانه يعطي نفس النتائج أو نتائج متقاربة .

ويقاس الثبات إحصائيا من خلال حساب معامل الارتباط بين الدرجات التي حصل عليها الأفراد في المرة الأولى وبين نتائج الاختبار في المرة الثانية وهو نسبة التباين الحقيقي الداخل في تباين الدرجات التجريبية. ويمكن إن نستدل من صدق الاختبار على انه ثابت ، في حين إن الاختبار الثابت ليس بالضرورة إن يكون صادقا

ومن المهم أن نعي بان مفهوم الثبات هو مؤشر على استقرار النتائج على أداة القياس ، ومن هنا فان الثبات ليس للأداة وإنما للنتائج على الأداة ويتم الحكم على ثبت أداة القياس يتم من خلال استقرار النتائج عليها ويتم الاستدلال على الثبات من خلال ثلاثة أساليب هي تطبيق وإعادة تطبيق الاختبار ، وتطبيق نماذج متكافئة وحساب الاتساق الداخلي للأداة أو الاختبار ، علما بان هذه الأساليب قد تؤدي إلى تقديرات مختلفة لقيمة معامل الثبات ، وسيتم تناول هذه الأساليب كل على حدة، حيث تستخدم فكرة الارتباط للتعبير عن قيمة معامل الثبات

طرق حساب الثبات :

من المهم أن نشير إلى مصادر الخطأ التي تؤدي إلى وجود ما يسمى بدرجات الخطأ ، حيث تكمن مصادر الأخطاء في المفحوص وكذلك في أداة القياس كالتحيز وعدم وضوح الفقرات أو طول الاختبار أو عدم ملائمة نوع الفقرات .. . الخ وكذلك في الإجراءات التي تتم فيها عملية القياس كالمتطلبات المادية كالتهوية

والإضاءة وازدحام المكان والمراقبين وتوقيت الاختبار والعلم ، ويتم تقدير الثبات من خلال مجموعة من الطرق سنتحدث عن كل واحدة منها على النحو التالي:

أولا : طريقة تطبيق وإعادة تطبيق الاختبار Test-Re-test.

يشير اسم الطريقة إلى إجراءاتها ، حيث تطبيق الاختبار المنوي التوصل إلى مؤشر على ثباته مرتين بفاصل زمني م (١ - ١٤) يوم على نفس المفحوصين ويتمثل معامل ثبات الاختبار في معامل الارتباط بين الدرجات في مرتي التطبيق ، وتتوقف قيمة معامل الثبات على العلاقة بين الدرجات في مرتي التطبيق ، فإذا حصل كل طالب على نفس الدرجة في التطبيقين فان الثبات سيكون عاليا ، أو إذا ارتبطت الدرجات في مرتي التطبيق ارتباطا خطيا فان ثبات الاختبار يكون عاليا، ويمكن أن يكون الثبات عاليا ، وكذلك إذا كان وضع الطلبة في مرتي التطبيق متشابه أي حافظ كل طالب على ترتيبه في مرتي التطبيق لكن هذه الطريقة محفوفة ببعض المخاطر التي قد تؤثر في قيمة معامل الثبات زيادة أو نقصانا

في هذه الطريقة يتم تطبيق الاختبار مرتين على نفس أفراد العينة(المفحوصين) تحت ظروف متشابهة قدر الإمكان وبفاصل زمني يقدر عادة من (1 - ١٤) يوم . ومن ثم حساب قيمة معامل الارتباط (بيرسون) بين الدرجات في مرتي التطبيق ويشير معامل الارتباط هنا الى ثبات الاختبار وعبر عن استقرار الأداء في مرتي التطبيق وهذا يعبر عن استقرار النتائج ولذلك يسمى معامل الثبات بمعامل الاستقرار. ويعد هذا النوع من ابسط الطرق المتبعة لتعيين معامل الثبات ، ويصلح في حساب معامل الثبات للاختبارات غير المحددة بوقت معين مثل الاختبارات النهائية ، ويفضل في هذا النوع ألا يكتفي بحساب الثبات على مدى فترة زمنية واحدة بل أكثر من فترة زمنية ثم أجراء معامل الارتباط بين كل فترة زمنية وأخرى ثم نأخذ المتوسط لمعاملات الارتباط المحسوبة ، هذا وتختلف المدة أو الفترة الزمنية بين التطبيق الأول والثاني .ففي اختبارات (الورقة والقلم) يفضل آن لا تقل الفترة الزمنية عن أسبوعين ويفضل تكرار التطبيق مرة أخرى وتصلح هذه الطريقة للاختبارات والمقاييس المقننة نظرا لتشابه الظروف والإجراءات المتعلقة بالتطبيق أكثر من غيرها.. أما الاختبارات البدنية يفضل

أن تكون الفترة الزمنية قريبة حيث لا يتأثر أداء الفرد بالتدريب ويذكر أن هناك بعض العوامل التي تؤثر في أداء الفرد وبالتالي في قيمة معامل الثبات وبالتالي في دقة الاختبار ومن هذه الأخطاء ما يلي :

- التذبذب العشوائي في أداء الفرد الواحد في المرتين حيث قد يحدث اختلاف في الأداء يؤدي لتذبذبه فيكون غير نمطي
- قد يتعرض الفرد لبعض التغيرات خلال الفترة الزمنية بين التطبيقين (نفسية، صحية، اجتماعية، ..الخ) مما يجعله قلقا أو متعجلا في أدائه مرة أو متأنيا مرة أخرى.
- تأثر الدرجات في التطبيق الثاني بالدرجات في التطبيق الأول وهو ما يسمى بأثر الحمل(Carry-Over effect) ويتعلق بتذكر المفحوصين لاجاباتهم في التطبيق الأول أو الألفة مع الاختبار بسبب تطبيقه مرتين فقد يؤدي إلى زيادة الثبات أكثر مما هو عليه.
- عند حدوث تغييرات في موقف الاختبار أثناء إعادة الاختبار كالمكان أو التوقيت أو التعب أو الظروف الجوية أو الإضاءة أو تسلسل أجزاء الاختبار .
- عند تعرض السمة أو القدرة المقاسة للتغيرات نتيجة التمرين والتعلم .

لذلك لابد من الضبط الدقيق لموقف الاختبار حتى لا تؤثر مثل تلك العوامل على معامل الارتباط بين مرتي التطبيق وكذلك الاهتمام بدراسة طبيعة السمة او القدرة التي يقيسها الاختبار لأن تعدد الطرق جاء نتيجة لتعدد طبيعة السمات المقاسة، وهذا يعني ان طريقة ما قد تكون اكثر ملائمة من طريقة أخرى

ثانيا : طريقة الصور المتكافئة : Alternate-Forms

يستخدم مصمم الاختبار هنا صيغتين متوازيتين او متكافئتين للاختبار الذي يطبق على نفس المجموعة من الأفراد ثم حساب معامل الارتباط بين الدرجات على الصيغتين أو الصورتين. وتسمى أحيانا الاختبارات المتوازية وهي التي لها نفس المتوسط ونفس التباين والتي ترتبط فيما بينها بنفس القدر لكن هذا صعب من الناحية العملية

لذلك يستعاض عنها بالصور المكافئة أو المتبادلة، أي انه في هذه الطريقة يوجد اختبارين لنفس السمة ويوجد تطبيقين ومجموعة واحدة من المفحوصين ويمتاز هذه الطريقة بتوفير الوقت والجهد في التطبيق لكنها لا تصلح لجميع الاختبارات (كالاستبيان والمقابلة الشخصية وغيرها). ويراعى في هذه الطريقة تكافؤ فقرات الاختبار في الصورتين من حيث الصعوبة والتمييز ومدى تمثيل العبارات للسمة المقاسة وتشابه المحتوى وتحوي نفس العدد من الأسئلة وأسلوب صياغة الفقرات بالإضافة إلى تكافؤ تعليمات الاختبار في الصورتين وعدد العبارات ، حيث ان عدم تكافؤ صورتي الاختبار يؤدي إلى عدم التقدير الدقيق لثبات الاختبار، وكذلك الفترة الزمنية المخصصة لتطبيق الاختبار وكذلك إجراءات التطبيق كما لا بد من الانتباه إلى الفترة الزمنية بين التطبيقين حيث أنها إذا قصرت أو طالت يصبح ثبات الاختبار مهددا بنفس العوامل التي ذكرناها في طريقة تطبيق وإعادة تطبيق الاختبار، ومن المهم أن نشير إلى أن أفضل تقدير لمعامل الثبات بين الدرجات الظاهرية في هذه الطريقة يتحقق إذا كانت الاختبارات متوازية أو ترتبط بعلاقة خطية شريطة أن لا يؤثر عامل انتقال الأثر مؤثرا في الثبات

مفهوم التوازي paralleled

يتعلق مفهوم التوازي بالاختبار حيث يشير هذا المفهوم الى توفر مجموعة من الشروط في نموذجين لاختبار يقيس نفس السمة ، وتتلخص هذه الشروط بان الاختبارين تم بناءهما بنفس المواصفات وهذا يعني ان درجات المفحوصين على النموذج الأول من الاختبار تساوي درجاتهم على النموذج الثاني من الاختبار أي ان:

(م ١ = م ٢) وبالطبع هذا يعني أن (م ح ١ = م ح ٢)

كما أن الوسط ت الحسابي للدرجات على النموذج الأول يساوي المتوسط الحسابي للدرجات على النموذج الثاني وكذلك تباين درجات المفحوصين على النموذج الأول يساوي تباين درجاتهم على النموذج الثاني، كما يتطلب مفهوم

التوازي أن يكون معامل الارتباط بين الدرجات على النموذج الأول مع متغير (س) هو نفس معامل الارتباط بين النموذج الثاني مع نفس المتغير (س) كما أنه من الضروري أن يكون التباين المشترك بين النماذج المتوازية متساويا.فإذا توافرت تلك الخصائص في اختبارين أو أكثر فانه يمكننا حينها أن نقول بان النماذج متوازية (parallelled tests) ، لكن الحصول على هذه المتطلبات بسبب صعوبة الحصول على الدرجة الحقيقية ، وللتسهيل يتم اللجوء عادة إلى نوع آخر من الاختبارات والتي تتحرر من بعض شروط أو متطلبات التوازي وتسمى هذه الاختبارات بالاختبارات المتكافئة (Equivalent tests) ومن شروط ومتطلبات التكافؤ أن يكون الاختبارات تم تصميمها بنفس المواصفات ، وتساوي التباينات بين الدرجات على كلا الاختبارين وكذلك التباينات المشركة والتباينات الداخلية بين الدرجات ، لكن ليس من الضروري أن تتساوى الدرجات الحقيقية على كلا الاختبارين

ثالثا: طريقة التجزئة النصفية Split Halves

في هذه الطريقة يتم تطبيق الاختبار على المفحوصين وبعد الانتهاء من التطبيق يتم تقسيم الاختبار إلى نصفين ثم يحسب معامل ارتباط بيرسون بين الدرجات على النصفين بمعنى أننا نكون قسمنا الاختبار إلى جزأين ، لكن من المهم أن ندرك وننوه إلى أن معامل الثبات المحسوب بهذه الطريقة يعبر عن ثبات نصف الاختبار ، ولذلك يتوقع أن يكون ثبات الاختبار ككل أعلى من ثبات نصفه بسبب زيادة طول الاختبار أي عدد فقراته حيث أن طول الاختبار يضمن تمثيلا أكثر للمحتوى مما يعني صدق محتوى أعلى ، وهذا يزيد من دقة الاختبار الذي يعبر عنها عادة بالثبات ، وجدير التذكير بأن ثبات التجزئة النصفية يفترض تكافؤ نصفي الاختبار ، وحيث أن هذا الافتراض صعب التحقق فانه لا بد من تصحيح معامل الثبات المحسوب لنصف الاختبار، وذلك من خلال معادلة سبيرمان براون على اعتبار أن النصفين متكافئين ، أما في حال عدم تكافؤ النصفين فإننا نلجأ لاستخدام معادلة كرونباخ الفا ، أما كيفية التقسيم فقد يكون بأخذ الفقرات الفردية في جزأ والفقرات الزوجية في الجزء الثاني أو اخذ النصف الأول كجزء والنصف الثاني كجزء ثانيا و توزيع الفقرات حسب

خصائصها السيكومترية (الصعوبة والتمييز) بحيث يكون الجزاين متكافئين مـن حيـث الصعوبة والتميز وعلى الرغم من اختلاف هذه الطرق في أسلوب تنصيـف الاختبـار لكنها تتفق في كيفية حساب معامل الارتباط أما المعادلات المستخدمة فهي :

في حال تكافؤ نصـفي الاختبـار نستخدم صـورة مـن معادلـة سـبيرمان بـراون لاختبار مكون من جزاين (نصفين) وهي كما تظهر في المعادلة رقم (٣٢)

$$\text{أ ث} = \frac{٢ \times \text{م}_{\text{ر}}\ ١.٢}{١ + (\text{م}_{\text{ر}}\ ١.٢)} \qquad (٣٢)$$

حيث.

م ث: معامل ثبات الاختبار ككل.

م ر ١.٢: معامل الارتباط بين نصفي الاختبار ورقم (٢) هو عدد الاختبارات (النصفين) لاحظ أن هذه الصورة مشتقة من المعادلة الأصلية لكن القائم فيها (١ + ٢-١) الذي يلغي (١+ ر١، ٢) على اعتبار أن ن=٢

في حال عدم توازي أو تكافؤ نصفي الاختبار أي في حال عدم تسـاوي التبـاين لدرجات المفحوصين على نصفي الاختبـار (غير متكـافئين) يمكـن استخدام معادلـة كرونباخ الفا كما في الصيغة رقم (٣٣)

$$\text{م ث} = \frac{٢\,(\text{ع}^{٢}\text{ه} - \text{ع}^{٢}\text{د}١ + \text{ع}^{٢}\text{د}٢)}{\text{ع}^{٢}\text{ه}} \qquad (٣٢)$$

حيث.

$\text{ع}^{٢}$ه : تباين درجات الاختبار ككل.

$\text{ع}^{٢}$د١: تباين درجات النصف الأول.

$\text{ع}^{٢}$د٢: تباين درجات النصف الثاني

كما ويمكن حساب التباين المشترك بين نصفي الاختبار باعتباره يمثل بسط المعادلة () وذلك على النحو المبين في المعادلة رقم (٣٤).

$$م_{ت} = \frac{2\,(ع^{2}\ ش_{1،2})}{ع^{2}\ ك} \quad \ldots\ldots\ldots\ldots\ldots(34)$$

حيث

$ع^{2}$ ك: التباين الكلي للدرجات المفحوصين لنصفي الاختبار معا.

$ع^{2}$ $ش_{1،2}$: التباين المشترك بين درجات المفحوصين على نصفي الاختبار.

ولا بد من معرفة أن قيمة معامل الثبات المحسوبة باستخدام كل من معادلتي سبيرمان براون وكرونباخ–الفا تزداد بزيادة قيمة معامل الارتباط بين نصف الاختبار والعكس صحيح ، ويكون الارتباط عاليا عندما يقيسان نفس السمة ، وان الارتباط بينهما عالي ، أي عندما يكون بينهما تباينا مشتركا وحينها يمكن حساب قيمة معامل الثبات بين نصفي الاختبار إذا استطعنا إيجاد التباين المشترك بين درجات المفحوصين على نصفي الاختبار من خلال المعادلة أعلاه ، وحيث أن معادلة سبيرمان براون تفترض تجانس التباين للاختبارين أي تساوي التباين للدرجات المفحوصين عل نصفي الاختبار ، لكن هذا الافتراض قد لا يتوفر أو يتحقق وهذا يؤدي إلى تقدير غير دقيق لثبات الاختبار ككل ، لذلك اقترح جتمان Guttmann معادلة تأخذ في الحسبان اختلاف تباين الدرجات على نصفي الاختبار وهي المعادلة التي تظهر في المعادلة رقم (٣٥)

$$ر_{س\,س} = 1 - \frac{ع^{2}_{1} - ع^{2}_{2}}{ع^{2}_{س}} \quad \ldots\ \ldots\ \ldots(35)$$

$ر_{س\,س}$: معامل ثبات الاختبار ككل.



$ع^2$ ١ : تباين النصف الثاني من الاختبار .

$ع^2$ ٢ : تباين النصف الثاني من الاختبار .

$ع^2$ س: التباين الكلي للاختبار .

كما اقترح رولون (Rulon) معادلة أخرى تعتمد على تباين الفروق بين درجات المفحوصين على نصفي الاختبار والتباين الكلي لاختبار وتظهر كما هي في المعادلة رقم (٣٦)

$$ر_{س س} = ١ - \frac{ع^2_{د}}{ع^2_{س}} \quad ....(٣٦)$$

حيث.

$ر_{س س}$. معامل ثبات الاختبار ككل.

$ع^2_{د}$:تباين الفروق بين الدرجات على نصفي الاختبار .

$ع^2_{س}$: التباين الكلي للاختبار .

ومن المهم ان نعلم أن قيمتي معامل الارتباط المحسوب بمعادلتي سبيرمان براون وكرونباخ- ألفا يمكن تتساويان إذا تساوت قيم تباين كل من درجات المفحوصين على نصفي الاختبار ، لكن إذا كان النصفين غير متكافئين رغم تساوي قيم التباين لكل منهما فان ذلك سيؤدي إلى تخفيض قيمة معامل الثبات، أما إذا توفرت شروط التكافؤ وتساوت قيم التباين فستتساوى قيم معاملات الثبات بالطريقتين. كما ويمكن تطبيق المعادلة السابقة على الاختبار العادي وذلك بحساب المتوسط الحسابي لقيم معاملات الارتباط الداخلية بين الفقرات وتعويضها بدل قيمة $م ر_{١٢}$ الواردة في المعادلة أعلاه على اعتبار ان الارتباط بين نصفي الاختبار هو ارتباط واحد (بين جزأين أو متغيرين).

رابعا: طريقة الاتساق الداخلي Internal Consistency

وتعتمد على استخدام كل من تباين فقرات الاختبار وتباين الاختبار ككل لتقدير الثبات وهي من أكثر طرق تعيين الثبات شيوعا ، حيث يطبق الباحث الاختبار أو الاستبيان ....الخ مرة واحدة ، أي يعطى الفرد درجة واحدة لجميع المفحوصين بحيث يطبق الاختبار مرة واحدة وتستخدم الدرجات على فقرات الاختبار أي الدرجات الجزئية أو الدرجات الكلية لتقدير ثبات الاختبار ، ولهذه الطريقة صيغة عامة لحساب معامل الثبات وتظهر الصيغة العامة لحساب الثبات بغض النظر عن أنواع الفقرات كما في المعادلة رقم (٣٧)

$$م ث = ن/(ن-١) \frac{ع^2 ك - مجموع ع^2 ف}{ع^2 س} \quad ...... (٣٧)$$

حيث.

ن: عدد الفقرات المكونة للاختبار.

$ع^2$ ك : تباين الدرجات على الاختبار ككل.

مجموع $ع^2$ف: مجموع تباين كل فقرة من فقرات الاختبار.

إلا أن طبيعة الاختبار والغرض منه ونوع الفقرات وكذلك المعلومات المتوفرة عن إحصائيات الاختبار دفعت علماء القياس إلى اشتقاق عدة معادلات تلعب ذلك باستخدام عدة طرق حيث تعتمد هذه الطرق على نوع فقرات الاختبار وطبيعة الغرض من الاختبار ومن هذه الطرق ما يلي:

الحالة العامة طريقة سبيرمان – براون :

تقوم معادلة سبيرمان براون على أن كل فقرة من فقرات الاختبار تعتبر مكونا من مكونات الاختبار ولها تباينها الخاص بها بمعنى أن المفحوصين يتباينون في الإجابة، ويتم حساب قيمة معامل الثبات للاختبار من خلال فكرة الارتباطات الداخلية للفقرات حيث ستكون هذه الارتباطات حسب عدد الفقرات حيث تكون عدد هذه الارتباطات حسب المعادلة التالية

ن(ن-١)/ ٢ فإذا كان عدد الفقرات (٢٠) فسيكون عدد الارتباطات الداخلية هو (٢٠ ×١٩)/ ٢ =(١٩٠) ارتباط داخلي

وأما الصيغة العامة لمعدلة معامل الثبات حسب طريقة سبيرمان براون والتي تشتق منها صيغ أخرى لحساب معامل الثبات فهي كما في المعادلة رقم(٣٨).

$$ر_{س س} = \frac{ن\,(ر^{-}\,ف)}{1 + (ن_1) \times ر^{-}\,ف} \qquad \text{.. .... ..} \quad (٣٨)$$

حيث:

ر س س: معامل ثبات الاختبار

ن: عدد الفقرات

ر¯ : المتوسط الحسابي لمعاملات الارتباط الداخلية بين الفقرات...............

إن استخدام معادلة سبيرمان براون يساعدنا في تحديد معامل ثبات الاختبار إذا تم تجزئته إلى نصفين(ثبات التجزئة النصفية) أو إلى عدة أجزاء ، كما ويمكن من خلالها تحديد عدد الفقرات المناسب للحصول على قيمة محددة لمعامل الثبات ، وحيث أن حساب معاملات الارتباط الداخلية بين الفقرات المكونة لأي اختبار سيحتاج للوقت والجهد الكبيرين خاصة إذا كان عدد الفقرات كبيرا فقد تم اللجوء إلى فكرة تباين الفقرات المكونة للاختبار ولذلك تم اشتقاق أكثر من صيغة لهذه المعادلة تفرضها طبيعة الفقرات والمعلومات المتوفرة عن الاختبار ، أما هذه الصيغ فهي :

دار جرير للنشر والتوزيع

معادلة كرونباخ الفا :

وتمثل الصيغة العامة لحساب ثبات الاختبار من خلال التجانس الداخلي وتتمثل في اللجوء إلى عدد الفقرات ومجموع تباينات الفقرات المكونة للاختبار وكذلك تباين الاختبار ككل وهي كما في الصيغة الواردة في المعادلة (٣٩)

$$ر_\alpha = ن/ن_١ \left(١ - \frac{مج ع^2 ف}{ع^2 س} \right) \quad (٣٩)$$

حيث:

رα: معامل الثبات بطريقة كرونباخ الفا.

ن:عدد فقرات الاختبار.

مجموع ع٢ف: مجموع تباين جميع الفقرات المكونة للاختبار كل على حدة.

ع٢س:التباين الكلي للاختبار.

وتستخدم هذه المعادلة لحساب ثبات الاختبار بغض النظر عن نوع الفقرات التي يتكون منها الاختبار. أننا لا بد أن ننتبه إلى قضية مهمة وهي انه إذا كانت تباينات الفقرات المكونة للاختبار أو الأجزاء المكونة للاختبار متكافئة وكانت تقيس سمة واحدة فان أفضل تقدير لمعامل الثبات سيكون باستخدام معادلة كرونباخ ألفا ، أما إذا لم تتوفر ظروف التكافؤ للفقرات أو للأجزاء فان معامل الثبات حسب هذه المعادلة سيكون اقل من قيمة معامل الثبات الحقيقية.

معادلة كودر-ريتشاردسون (٢٠)

وسميت بهذا الاسم لانها تمثل المعادلة رقم (٢٠) لكل من كودر وريتشارسون وهي صورة مشتقة من معادلة كرونباخ الفا ولكن تستخدم في لحساب ثبات الاختبار المكون من فقرات الاختبار من متعدد من خلال الاستفادة من معامل الصعوبة والسهولة لكل فقرة من فقرات الاختبار ، حيث يتم حساب تباين الفقرات من خلال

حاصل ضرب معامل الصعوبة في معامل السهولة وتظهر هذه الصيغة في المعادلة رقم (٤٠)

$$KR\text{-}20 = \frac{ن}{ن-1}\left| 1 - \frac{مجـ ص(1-ص)}{ع^2_س} \right|$$
(٤٠)...............................

حيث

ث KR-2. :معامل الثبات المحسوب بمعادلة كورد –ريتشارسون ٢..

ن:عدد الفقرات.

ص⁻ :معامل صعوبة الفقرة.

(١ – ص) :معامل السهولة للفقرة (متممة معامل الصعوبة).

ع٢س: التباين الكلي للاختبار ككل.

معادلة كودر – ريتشاردون ٢١.

وسميت بهذا الاسم لنفس السبب الوارد في المعادلة رقم (٢.) ، وتستخدم هذه المعادلة لحساب معامل ثبات الاختبار المكون من فقرات الاختيار من متعدد لكنها تفترض أن جميع الفقرات متساوية معامل الصعوبة ، وهذا صعب من الناحية العملية ، ومن هنا فهي قليلة الاستخدام وتعتمد هذه المعادلة على استخدام كل من المتوسط الحسابي والتباين لدرجات المفحوصين على الاختبار ككل وتظهر الصيغة المستخدمة في المعادلة رقم(٤١)

$$KR\text{-}21 = \frac{ن}{ن-1}\left| 1 - \frac{م(ن-م)}{ن ع^2_س} \right|$$
(٤١)...............................

حيث:

ث KR-21 : معامل الثبات المحسوب بمعادلة كورد –ريتشارسون ٢٠.

ن: عدد الفقرات.

س ¯ : المتوسط الحسابي لدرجات المفحوصين على الاختبار.

ع٢س: التباين الكلي للاختبار ككل.

وبالنسبة للاختلافات التي قد تظهر بين قيمة معامل الثبات المحسوب بالمعادلتين (٢٠ ، ٢١) فانه لا يوجد اختلاف إذا كانت معاملات الصعوبة متساوية للفقرات ، أما الثبات المحسوب بمعادلة كرونباخ الفا ومعادلة كودر – ريتشاردسون (٢٠) فسيعبر فعلا عن ثبات الاختبار إذا كانت الفقرات متكافئة ، وبالنسبة للثبات المحسوب باستخدام معادلة (٢١) فسيعبر عن ثبات الاختبار فعلا إذا كانت صعوبة الفقرات متساوية ومتكافئة على اعتبار أن كل فقرة منها تعتبر اختبارا جزئيا من الاختبار ككل ، وبالنسبة لقيمة معامل الثبات باستخدام المعادلات الثلاث (الفا، ٢٠ ، ٢١) ستكون اكبر ما يمكن إذا كانت الارتباطات الداخلية بين مكونات الدرجات الحقيقية للمفحوصين عالية وهذا لا يتحقق إلا إذا كانت الفقرات تقيس سمة واحدة.

كما ويمكن إيجاد معامل ثبات الاختبار من خلال فكرة تحليل التباين واستخدام فكرة مصادر التباين الذي يحدده نوع تحليل التباين والذي يحدده بالطبع عدد المتغيرات المستقلة ، حيث يتم حساب تباين درجات المفحوصين لجميع الفقرات واستخدام فكرة التباين الحقيقي والتباين الكلي وتباين الخطأ حيث يمثل معامل الثبات نسبة التباين الحقيقي إلى التباين الكلي، وكما هو معروف فان التباين الحقيقي هو الفرق بين التباين الكلي وتباين الخطأ اي أن

رث = ع٢ح/ع٢ه

حيث أن ع٢ح= ع٢ه – ع٢ع.

اي أن رس = ع٢ه – ع٢ع /ع٢ه .

حيث : ع٢ح التباين الحقيقي، ع٢ه التباين الكلي، ع٢ع:تباين الخطأ.

ثبات الاختبار محكي المرجع

من المعروف بأن الاختبار قد يكون معياري المرجع أو محكي المرجع ، حيث أن الفرق بين الاختبارين يكمن في طريقة تفسير النتائج إذ تفسر نتائج الاختبار معياري المرجع في ضوء أداء المجموعة في حين تفسر نتائج الاختبار محكي المرجع في ضوء مستوى معين أو درجة قطع محددة ، ويتوقع أن يكون معامل الثبات للاختبار المحكي لوجود درجة قطع محددة الأمر الذي يؤدي إلى ضيق مدى الدرجات ، لذا يمكن اللجوء إلى عملية تصحيح لمعامل الثبات من خلال استخدام بعض المعادلات التي تعالج خصوصية الاختبار محكي المرجع ومن هذه المعادلات :

معادلة ليفنجستون

وقد اقترحها ليفنجستون (Livingston,1972) ولذلك سميت باسمه حيث يتم اللجوء إليها لتصحيح قيمة معامل الثبات المحسوب بالطرق سالفة الذكر، أما الصيغة العامة لهذه المعادلة فتظهر في المعادلة رقم (٤٢)

$$\text{ر}'_{\text{ع}} = \frac{\text{ر}_{\text{ع}} \times \text{ع}^2_{\text{س}} + (\bar{\text{س}} - \text{م})^2}{\text{ع}^2_{\text{س}} + (\bar{\text{س}} - \text{م})^2} \quad \text{.........................} (٤٢)$$

حيث

ر'ع: معامل الثبات المصحح (لاختبار المحكي المرجع).

رع: معامل الثبات المحسوب بالطريقة العادية.

$\text{ع}^2_{\text{س}}$: تباين الدرجات على الاختبار.

($\bar{\text{س}}$): المتوسط الحسابي لدرجات الطلبة على الاختبار.

م : درجة القطع المعتمدة للاختبار.

وبما أن الاختبار محكي المرجع يهتم بتصنيف المفحوصين إلى مجموعتين من خلال فكرة درجة القطع (Cut Score) إحداهما متمكنة والأخرى غير متمكنة فإن قيمة معامل الثبات المحسوبة سوف تتأثر بقيمة درجة القطع المحددة للنجاح على

الاختبار ، فاذا تغيرت قيمتها سوف تأثر قيمة معامل الثبات ، وكذلك في حال عدم ملائمة درجة القطع المعتمدة ، ومن المهم أن ندرك ان الاختلاف بين هذه الصيغة وصيغ معادلات الثبات التقليدية انها تأخذ بالحسبان الفرق بين متوسط درجات المفحوصين ودرجة القطع المعتمدة ، وتعتمد على تباين الدرجات الحقيقية والملاحظة ولذلك يمكن كتابتها على صيغة أخرى كم تظهر في الصيغة رقم().

$ع^{2}_{م} + (\bar{س} - د\,ق)^{2} \;/\; ع^{2}_{ح} + (\bar{س} - د\,ق)^{2}$. حيث.

$ع^{2}_{م}$: تباين درجات المفحوصين حول درجة القطع.

$\bar{س}$: المتوسط الحسابي للدرجات الطلبة.

د ق : درجة القطع المعتمدة للنجاح على الاختبار.

$ع^{2}_{ح}$: تباين الدرجات الحقيقية للمفحوصين حول درجة القطع.

وقد بدأ الاهتمام منذ محاولات ليفنجستون بثبات الاختبار محكي المرجع فتوالت عدة محاولات لمعالجة مشكلة خاصة بقيم معاملات الثبات حسب معادلته حيث أنها لا تتصف بخاصية عدم التغير ولذلك من الصعب إجراء التحويلات الخطية عليها.

معادلة كابا ، سوامنيثان، هامبيلتون، الجانيا

وقد جاءت محاولاتهم هنا لإيجاد مقياس إحصائي يعبر عن اتساق التصنيف للمفحوصين في مرتي التطبيق للاختبار عند استخدام طريقة تطبيق وإعادة تطبيق الاختبار، مع الافتراض باستبعاد عوامل الصدفة في التصنيف ، وقد توصلوا إلى الصورة الرياضية التالية لهذا المقياس والذي سمي بمعامل كابا كما في المعادلة رقم (٤٣)

$$م_{ه} = \frac{ن - ن_{م}}{1 - ن_{م}} \quad ....(٤٣)$$

حيث:

م ه : معامل الارتباط المحسوب بطريقة كابا.

ن: نسبة الاتفاق الملاحظ لتصنيفات المفحوصين.

ن م: نسبة الاتفاق المتوقع لتصنيفات المفحوصين.

مع الانتباه إلى أن

(ف) تمثل مجموع نسب المفحوصين المصنفين في نفس المستوى (س) في مرتي التطبيق وأن هناك عدد ك من المجموعات المصنفة.

(فم) تمثل مجموع النسب للمفحوصين المصنفين في المستوى(س) في مرتي التطبيق اي أن (ف م) ستكون من نسبة المصنفين في المستوى (س) في مرة التطبيق الأولى ولتكن (ن م١) ونسبة المفحوصين المصنفين في المستوى (س) في مرة التطبيق الثانية (ن م٢) ولتكن.وهذا يشير إلى أن قيمة معامل الثبات المحسوبة بهذه الطريقة تعتمد على نسب الاتفاق في تصنيف الفرد في مرتي التطبيق اي أن الاتفاق في التصيمين يزيد من ثبات الاختبار وهذا منطق لان يحقق معنى الاستقرار للنتائج في مرتي التطبيقين، كما أن درجة القطع المعتمدة تؤثر في قيمة معامل الثبات ، إضافة إلى عدد ونوعية المفحوصين لأننا نتحدث عن مجموع نسب التصنيف التي تعتمد على عدد أزواج التصنيف لكل مفحوص في مرتي التطبيق، وكذلك يتأثر معامل الثبات بهذه الطريقة بصعوبة الفقرات وعددها ، وسيتضح ذلك في الأمثلة التوضيحية التي سنوردها لاحقا.

معادلة كابا المعدلة

اهتمت المعادلة السابقة لكابا بثبات الاختبار المطبق مرتين أو ذو الصورتين المتكافئتين وحاول هاينا Huynh(١٩٧٤) الاهتمام بثبات الاختبار المحكي المرجع الذي يطبق لمرة واحدة ، وقد استند إلى معادلة كابا بحيث يستخرج معامل لإيجاد المعامل من نتائج التطبيق لمرة واحدة ، وافترض لذلك أن يكون الاختبار مكونا من مجموعة من الفقرات المختارة عشوائيا من نطاق (مجتمع) شامل للسمة المقاسة بحيث يكون معرفا ومحددا بشكل جيد ولعل هذا المعامل يناسب الاختبارات المحكية مرجعية النطاق (Domain-Referenced Test) إضافة إلى أن تتوزع الدرجات الحقيقية لكل مفحوص على حدة شكل التوزيع بيتا ، وكذلك أن تكون المفردات ثنائية الدرجة(صفر، ١) ، ومن الانتقادات التي وجهت لهذا المعامل انه صعب التحقق لافتراضاته إضافة إلى كثرة العمليات الحسابية التي يحتاجها والتي لا مجال للحديث عنها في هذا الكتاب.

معادلة سابكوفياك

كان من بين افتراضات معامل كابا أن يكون عدد فقرات الاختبار كبيرا ، حيث تنعدم فائدة معامل كابا في حال كانت الفقرات المكونة للاختبار قليلة ، لذلك حاول Subkoviak (١٩٧٦) إيجاد معامل الاتفاق اعتمادا على فكرة الاحتمالات ، حيث اقترح استخدام احتمال تصنيف كل مفحوص إلى صنفين (متمكن، غير متمكن) حسب درجة قطع الاختبار ولذلك سمي مؤشر الثبات المحسوب بهذه المعادلة بمعامل الاتفاق ، حيث يتم تقدير احتمال تصنيف كل مفحوص تصنيفا صائبا والقيام بجمع هذه الاحتمالات وإيجاد قيمة المتوسط الحسابي لها بحيث يمثل هذا المتوسط معامل اتفاق خاص بالمفحوصين الذين يطبق عليهم الاختبار ويمكن إيجاد هذا المعامل من خلال المعادلة رقم(٤٤)

ل د = مجموع ل د(س) / ن . . . . (٤٤)

حيث .

ل د :معامل الاتفاق.

ل د(س):احتمال اتساق تصنيف المفحوص تصنيفا صائبا فيما لو طبق عليه اختبارين.

ن: عدد المفحوصين.

يذكر أن هناك حالتي تصنيف الأولى تكون درجتي الفرد أكبر أو تساوي درجة القطع والثانية أن تكون درجة الفرد اقل من درجة القطع لا ن تعريف سابكوفياك لمعامل الاتفاق هو احتمال التصنيف المتسق (متقن/ متقن أو غير متقن/ غير متقن) فيما لو خضع لاختبارين متوازيين ومحكي المرجع وبطبيعة الحال يتطلب ذلك ان تكون الفقرات ثنائية التصحيح(صفر، ١) وان تكون صعوبة الفقرات متساوية تقريبا في الصعوبة ، كما أن توزيع درجات أي مفحوص حسب فقرات كلا الاختبارين المفترضين سيكون ذي حدين ، وإذا لم تكن الفقرات متساوية في الصعوبة فإننا يمكن أن نستخدم نموذج ذي الحدين لتقدير مناسب لاحتمالات التصنيف، ولذلك فان هذه

الطريقة يمكن استخدامها إذا كان هناك أكثر من درجة قطع واحدة للاختبار. ولا يتسع المكان للحديث عن تفصيلات نموذج ذي الحدين.

وقد توالت محاولات عديدة نتيجة للاهتمام بتقدير ثبات الاختبارات محكية المرجع حيث برزت نظرية التعميم والتي وجهت كثيرا من جهود المهتمين في القياس بهذا المجال مثل معامل الاعتمادية لبرنان وكين.وفي النهاية فانه لا بد من التأكيد على أن أية محاولة لدراسة ثبات الاختبار بغض النظر عن فيما إذا كان معياري أو محكي المرجع لا بد أن ينطلق من المعنى العام للثبات وهو دقة تقدير قدرة المفحوص إضافة إلى علاقته بمفهوم صدق الاختبار.

العوامل المؤثرة على الثبات :

يعبر ثبات الاختبار عن دقة النتائج على الاختبار وكلما كان الثبات عاليا دل ذلك على انخفاض الفرق بين الدرجات الحقيقية والدرجات الملاحظة ، ولذلك فان كل ما يؤثر على دقة الاختبار يعتبر عاملا مؤثرا في الثبات ، ومن العوامل المؤثرة بثبات الاختبار ما يلي:

- تجانس مجموعات التطبيق حيث يقل معامل ثبات الاختبار لمجموعة متجانسة عن معامل ثبات نفس الاختبار لمجموعة أقل تجانسا.
- وقت الاختبار فزيادة وقت الاختبار تؤدي لزيادة ثبات الاختبار .
- وضوح الاختبار من حيث الصياغة والتعليمات وظروف التطبيق .
- التخمين حيث أن كلما خمن المفحوصين في الإجابة يقل ثبات الاختبار فقد تختلف الإجابة بالتخمين في مرتي التطبيق.
- الفقرات الغامضة والعاطفية والطويلة تقلل من ثبات الاختبار
- طول الاختبار : بمعنى انه إذا زاد عدد فقرات الاختبار زاد معامل الثبات
- درجة القطع المعتمدة بالنسبة للاختبارات محكية المرجع .
- تجانس فقرات الاختبار : يزداد معامل الثبات كلما كانت الفقرات متجانسة
- صعوبة الفقرات : تزداد قيمة معامل الثبات عندما تكون الفقرات معتدلة الصعوبة

- تميز الفقرات : يرتفع معامل الثبات عندما تكون الفقرات ذات تميز عال
- تجانس المفحوصين . يزداد معامل الثبات كلما كانت عينة المفحوصين متباينة .
- يتأثر الثبات باختلاف ظروف تطبيق الاختبار
- تؤثر حالة الفرد النفسية والصحية على ثبات الاختبار ، فالمرض والتعب والتوتر قد يؤدي إلى انخفاض الثبات .
- يتأثر الثبات بموضوعية التصحيح ' فتباين التصحيح يؤدي إلى زيادة تباين الخطأ وبالتالي إلى انخفاض الثبات .

كيف يؤثر طول الاختبار على الثبات.

يقصد بطول الاختبار زيادة عدد فقراته حيث يفترض أن زيادة عدد الفقرات يتيح تغطية أكبر للمحتوى وهذا يحقق الصدق المتعلق بالمحتوى ، إضافة إلى أن الزيادة في عدد الفقرات يؤدي إلى زيادة تمثيل فقرات الاختبار لمجتمع الفقرات التي من الممكن أن تقيس السمة التي اعد الاختبار لقياسها ، وهذا بدورة يؤدي إلى تخفيض الخطأ العيني مما يزيد في دقة الاختبار، وبالنظر إلى معادلات معامل الثبات نجد أن عدد الفقرات يعتبر جزءا من المعادلة ولذلك فانه ومن الناحية الرياضية فان عدد الفقرات يؤثر في قيمة الثبات حيث أن زيادة عدد فقرات الاختبار يؤدي رياضيا إلى زيادة قيمة معامل الثبات للاختبار ، ولذلك فان حساب معامل ثبات الاختبار بطريقة التجزئة النصفية يتطلب تصحيح معامل الثبات لان الذي يتم حسابه هو ثبات احد نصفي الاختبار وذلك باستخدام معادلة سبيرمان براون لهذه الغاية ولذلك فمن المتوقع زيادة معامل ثبات الاختبار لنصف الاختبار عن معامل الاختبار ككل لاحظ أن.

م ث= ن×ر/ ١+(ن-١) ر

وباعتبار أن عدد الأجزاء (ن) يساوي (٢) فان م ث= ٢×ر/ ١+ ر لاحظ أن زيادة ن تؤدي إلى زيادة معامل الثبات ولذلك يمكننا حساب عدد الفقرات التي توفر مستوى الثبات المرغوب ، أو الزيادة المطلوبة التي تحقق زيادة في قيمة معامل الثبات ، فإذا كان لدينا اختبار تم إيجاد ثبت نصفه الأول وكان يساوي(٠.٧) فان معامل ثبات



الاختبار ككل سيكون أعلى نتيجة لزيادة عدد فقراته (فقرات النصف الثاني) وسيكون طبقا لمعادلة التصحيح ٢(٠.٧)/ ١+ (٠.٧) = ١.٤/١.٧=٠.٨٢ لاحظ الزيادة كانت (٠.١٢) وهذه قيمة لها دلالتها في الدقة ، ولذلك إذا عرفنا قيمة معامل ثبات اختبار ما ورغبنا في زيادة ثباته يمكننا تحديد الفقرات التي نحتاجها لزيادة طول الاختبار لضمان الثبات الأعلى، وذلك من خلال المعادلة التالية ن = ن ٢/ن١ حيث تشير ن إلى نسبة زيادة الفقرات فإذا كان لدينا اختبار ما وكان عدد فقراته (٢) وحصلنا على قيمة معامل ثبات (٠. ٥) ورغبنا بزيادة معامل الثبات إلى (٠. ٦) فانه يلزمنا زيادة عدد فقراته من خلال المعادلة السابقة ن= ن٢/ن١ هذا يعني أنه لا بد من إيجاد نسبة الزيادة من خلال معادلة التصحيح حيث أن ر = ن ر/ ١+ (ن-١)ر = ٠.٦ =ن×٠.٥/ ١+ ن-١)× ٥ = وبحل المعادلة نجد أن (ن = ١.٥) وهنا يمكن أن نجد عدد الفقرات اللازم أي أن ١.٥ = ن٢/ ٢ ومنه ن٢ = ٣ أي انه حتى نحصل على معامل ثبات (٠.٦) لنفس الاختبار لا بد أن نزيد عليه (١.) فقرات تقيس نفس السمة التي يقيسها الاختبار، ولا بد من الاهتمام ببنائها واختيارها بحيث تؤدي إلى الزيادة في الدقة أي الصدق والثبات للاختبار.

مما سبق يتبين أن توفير الثبات العالي (المرغوب) للاختبار يتطلب تحقق العديد من الظروف والمتطلبات ، وهذا بالطبع يتطلب الوقت والجهد مع الوعي بطبيعة المتطلبات التي ذكرنا والاهم من ذلك الوعي بطبيعة السمة المقاسة ، وكذلك العلاقة بين مفهومي الصدق والثبات التي سنتكلم عنها بعد الحديث عن مفهوم الصدق في الفصل القادم.

أمثلة توضيحية لمفهوم الثبات

مثال(١) إذا كان معامل ثبات اختبار من نوع الاختبار من متعدد لقياس وحدة العمليات الحسابية في الرياضيات (٠.٦٠) ، احسب معامل ثباته إذا أصبح أربعة أضعاف عدد الفقرات التي كان عليها سابقا .

الحل: بتطبيق معادلة سبيرمان براون ينتج أن

رس س = ن (ر ف) / ١+(ن_١) × ر ف

٤×(٠.٦٠) / ١+(٤-١) ×(٠.٦٠) = ٢.٤/ ١+(٣)×(٠.٦٠) .

م ث= ٢.٤/ ٢.٨= ٠.٨٥.

مثال (٢) إذا كان معامل الثبات لاختبار ما هو (٠.٦) وكانت مدته (٢) دقيقة ، أردنا الحصول على معامل ثبات أعلى بان أصبح الاختبار ضعفي طوله الأصلي ، فما هو معامل الثبات الذي سنحصل عليه؟

الحل: بتطبيق معادلة سبيرمان براون لحصل على التالي

رس س = ن (ر ف) / ١+(ن_١) × ر ف

معامل الثبات = ٢×(٠.٦٠)/ ١ + (٢-١)(٠.٦٠).

= ١.٢/ ١.٦= ٠.٧٥ وهذا يعني انه لو زدنا عدد الفقرات بشكل أكثر لحصلنا على معامل ثبات أعلى.

مثال (٣) أختبار مكون من (٥٠) فقرة وكان معامل ثباته (٩) وأردنا اختصاره بحيث يصبح مكتون من (٤٠) فقرة احسب معامل ثباته بعد الاختصار من طوله؟

الحل : هناك بعض الصيغ المشتقة من معادلة سبيرمان براون تستخدم لحاسب الثبات في مثل هذه الحالة حيث يعتبر الاختبار هنا جزءاً حيث أننا أختصرنا منه (١٠).

فقرات أي ١/ ٥ من عدد الفقرات ولذلك فان معامل الثبات يمكن حسابه من المعادلة التالية

ر المصحح = ١/ ن × ر المحسوب / ١+١/ ن-١)* ر المحسوب.

= ١/ ٥× (.٩٠) / ١+(١/ ٥-١)* (.٩٠)

= .٢٠/ .٩٠/ ١+ (.٢٠-١)*(.٩٠)

= .١٨/ ١+(-.٨٠)*(.٩٠)

= .١٨/ ١+ (.٧٢)

= .١٨/ .٢٨ = .٦٤

مثال (٤) اختبار يتكون من (١) فقرات فإذا طبق وكان معامل ثباته بهذا الطول (.٨٠) فما هو الطول الذي يجب أن يكون عليه حتى يصبح معامل ثباته (.٩٠)

الحل: حتى نعرف طول الاختبار لا بد من إيجاد نسبة الزيادة في الفقرات ومن ثم نحصل على العدد المطلوب للزيادة ونضربها بطول الاختبار الأصلي وتحديد الطول الجديد للفقرات على النحو التالي. ومن الحالات التي يمكن اشتقاقها من معادلة سبيرمان براون صيغة يمكن من خلالها يمكن تحديد قيمة عدد الفقرات التي نحتاجها للحصول على معامل ثبات أعلى وذلك بإعادة ترتيب المعادلة على النحو التالي

نسبة الزيادة = (ر مرغوب) (١- ر محسوب) / ر محسوب (١- ر مرغوب)

= .٩٠× (١- .٨٠) / .٨٠× (١- .٩٠)

= .١٨/ .٠٨ = ٢.٢٥

وهنا ينتج أن

ن= الطول الجديد/ الطول الحالي = ٢٥ ٢= الطول الجديد/ ١

وبالضرب التبادلي ينتج أن الطول الجديد= ٢.٢٥×١٠ =٢٢.٥ ≈ ٢٣أي انه للحصول على معامل ثبات (.٩٠) لا بد من زيادة عدد الفقرات (٢٣) فقرة ليصبح الاختبار من (٣٣) فقرة.

مثال(٥): إذا كان معامل ثبات اختبار مكون من (٨٠) فقرة يساوي(.٩٠) إذا تم اختصار الاختبار إلى النصف (٤٠) فقرة احسب معامل ثباته بعد الاختصار

الحل: الذي تم هنا هو تنصيف الاختبار ولذلك يمكن استخدام صيغة سبيرمان براون والتي تختص بثبات التجزئة النصفية على النحو التالي

ر = ١/٥ر ÷ (١+(١/٥-١)ر)

ر = .٩٠ × .٨٠/.٤٠ ÷ (١+ (.٨٠/.٤٠-١)×.٩٠) = .٤٥/.٥٥ = .٨٢

مثال (٦) اختبار مكون من (٥٠) فقرة حسب معامل ثباته عندما طبق على عينة من المفحوصين المتقدمين للبعثات الدراسية وكان (. ٧٥) ، أراد مستخدمي الاختبار أن يكون الاختبار دقيقا بحيث يزيد معامل الثبات إلى (.٩٠) كون نتائج الاختبار سيتم تحديد نتائج ابتعاث مجموعة من المتفوقين ، ما هي نصيحتك للمستخدمين لزيادة دقة الاختبار؟

الحل: في هذه الحالة لا بد لمستخدم الاختبار أن يزيد من المصمم اختبار أطول بحيث يقيس نفس السمة أو السمات المحددة، ومن هنا سنقوم بتحديد نسبة الزيادة التي يتطلبها زيادة معامل الثبات من خلال العلاقة التالية.

نسبة الزيادة = (ررر) (١- ررر) / ررر (١- ررر)

= (.٩٠) ×(١- .٧٥)/(.٧٥ ×(١-.٩٠)

= .٢٢٥/ .٠٧٥= ٣ أي أننا نحتاج أن نضاعف الاختبار ثلاثة أضعاف وبالتالي فان الاختبار الذي يلبي رغبة المستخدم يجب أن يتكون من (١٥٠) فقرة.

مثال (٧) اختبار محكي المرجع يتكون من (١) فقرة من نوع الاختيار من متعدد تم بناؤه لاختيار عدد من المرشحين لدراسة الطيران فإذا تم تطبيقه على مجموعة من الطيارين في السنة الثانية للتدريب، وكان معمل ثباته (. ٨.) كم يصبح ثباته إذا تم تصحيح معامل الثبات المحسوب باستخدام معادلة ليفنجستون إذا كان الوسط الحسابي لدرجات المفحوصين (٨٥) وكان الانحراف المعياري (٤) إذا علمت أن درجة القطع المعتمدة (٧٨) ؟

الحل : من خلال معادلة ليفنجستون فان

ر= ر ×ع^{٢} م +(س¯ - د ق)^{٢} / ع^{٢} ح +(س¯ - د ق)^{٢}

ر= (.٨..) × (٢٥) +(٨٥ - ٧٨)^{٢} / ٢٥ + (٨٥ - ٧٨)^{٢} تذكر أن التباين هو(٥) ^{٢}

ر = ٢.+ ٤٩ / ٢٥+ ٤٩ = ٦٩/ ٧٤ = ..٩٣٢

مثال (٨) تم تطبيق اختبار محكي المرجع مكون من (٤) فقرات لتعيين (٣) عمداء الكليات في إحدى الجامعات من بين (٦) مرشحين. ولهذا الغرض طبق الاختبار مرتين وكانت نتائج المتقدمين على مرتي التطبيق كما هو في الجدول أدناه ما هو معامل الثبات المحسوب من هذه النتائج إذا اعتمدنا درجة قطع (٧٥)

الرقم	درجات التطبيق الأول					درجات التطبيق الثاني				
	١	٢	٣	٤	الكلي	١	٢	٣	٤	الكلي
١	١	صفر	صفر	صفر	١	١	١	صفر	صفر	٢
٢	١	١	١	١	٤	١	١	١	صفر	٣
٣	١	صفر	١	صفر	٢	١	١	١	صفر	٣
٤	صفر	صفر	صفر	صفر	صفر	صفر	١	صفر	صفر	١
٥	صفر	١	١	١	٣	١	١	١	صفر	٣
٦	صفر	١	١	صفر	٢	صفر	١	١	١	٣

تحسب المصفوفة لتصنيفات المفحوصين (العمداء المرشحين) من النتائج أعلاه

التطبيق الثاني		التطبيق الأول		المجموع
		ناجح	راسب	
	ناجح	٣٣	صفر	٣٣
	راسب	٣٣	٣٣	٦٦.
	المجموع	٦٦.	٣٣.	

لاحظ هنا ثبات الاختبار يشير مدى اتساق تصنيف كل مفحوص في مرتي التطبيق اعتمادا على درجة القطع (٧٥)، ومن خلال تطبيق المعادلة الخاصة بمعامل كابا أو ليفنجسترون المعدل ينتج ان

(ن ~ ن م)

م ك = ―――――

١ـ ن م

حيث أن (ن) = ٣٣..+٣٣.. = ٦٦..

وكذلك (ن م) = (٣٣) × (٦٧ .) (٦٧..) × (٣٣) = ٤٤.

م ه = ٦٦.. − ٤٤ / ١− ٤٤ = ٢٢./٥٦ . = ٣٩٣ .

لاحظ أن معامل الاتفاق في تصنيف المفحوصين في مرتي التطبيق منخفض نوعا ما وقد يعود ذلك إلى قلة عدد الفقرات وعلى العموم فان معامل الاتفاق هذا هو الذي سيرتكن إليه متخذ القرار في الاعتماد على هذا الاختبار في الاختيار من بين المرشحين لأنه يمثل مدى دقة الاختبار.

وعلى العموم لا بد للمهتمين أن يكونوا على وعي بمفهوم الثبات لان ذلك يمكنهم من تفسير الإجراءات والقرارات المتخذة اعتمادا على نتائج الاختبارات وكذلك اختيار أي الاختبار يمكن استخدامه مع الوعي بطبيعة السمة المقاسة والوعي كذلك بعلاقة الثبات بمفهوم الصدق الذي سيكون موضوع الفصل القادم

الفصل الثالث

مفهوم الصدق

يمثل مفهوم الصدق الخاصية الثانية لأدوات القياس ومنها الاختبارات سواء أكانت معيارية أم محكية المرجع ، ويكون الاختبار صادقا بالدرجة التي يحقق فيها الغرض الذي اعد من اجله ، وللصدق مجموعة من المظاهر أو المؤشرات، وحيث انه يتم الاعتماد على المعلومات التي يتم الحصول عليها نتيجة لاستخدام أدوات القياس فلا بد أن تتصف هذه الأدوات بالصدق إلى جانب الثبات الذي تحدثنا عنه سابقا، ويعرف صدق الاختبار بأنه **قدرة الاختبار على قياس السمة التي اعد وصمم لقياسها،** ولذلك إذا تم تطبيق اختبار ذكاء مثلا على مجموعة من المفحوصين وكان من بينهم بعض الأذكياء والبعض من متدنيي الذكاء وافرز الاختبار الطلبة حسب مستويات الذكاء التي يتصفون بها ، أي انه اظهر الاختلافات بين المفحوصين نسمي الاختبار صادقا ، وللصدق أهمية بالغة في بناء وتطوير أدوات القياس ومنها الاختبارات ، لان توفر الصدق للاختبار يعني توفر متطلب كبير للثبات، لكن الصدق كمفهوم له مجموعة من الأنماط ويتم التعبير عنه من خلال مفهوم معامل الارتباط بين مكونات (فقرات) الاختبار أو بين الاختبار ذاته واختبارات أخرى تقيس نفس السمة ، وقد يكون الاختبار صادقا من خلال طبيعة فقراته وتمثيلها لمحتوى السمة التي يقيسها، فلو عدنا إلى تعريف الاختبار بأنه **إجراء منظم لقياس عينة من السلوكات من خلال عينة من المثيرات (الفقرات)** ، فان توفر مفهوم التمثيل لعينة الفقرات أو عينة السلوكات المقاسة يعبر أيضا عن مفهوم الصدق، ولذلك يكون الاختبار صادقا إذا :

- إذا ميز بين المفحوصين المختلفين في الذكاء بالنسبة لاختبار ذكاء
- افرز الطلبة ذوي التحصيل العالي والتحصيل المتدني بالنسبة لاختبار تحصيلي.

- حصل الموظفين ذوي الأداء العالي على أعلى الدرجات.
- إذا استطاع التنبؤ بأداء الطلبة المبعوثين للدراسات العليا .
- إذا اظهر الفروق الفردية بين المفحوصين.

أنواع الصدق:

تتعدد أغراض القياس أو الاختبار وكذلك طبيعة السمة المقاسة وتبعا لذلك تعددت أشكال الصدق للاختبار ومن هنا فالصدق عدة أنواع هي على النحو التالي

أولا: صدق المحتوى:

يمثل صدق المحتوى أهم أنواع الصدق وهو ضروري لجميع أدوات القياس ولا سيما الاختبارات حيث لا بد أن تكون المعلومات التي نحصل عليها من الاختبار صادقة تعبر عن مقدار السمة المراد قياسها لدى المفحوص، والفكرة الرئيسية لصدق المحتوى أن يقيس الاختبار المحتوى (السمة) الذي اعد لقياسه ويتوفر صدق المحتوى بشكل عام من خلال جدول المواصفات الذي سنتحدث عنه فيما بعد ، وبطبيعة الحال فان هذا الصدق يتوفر من خلال مجموعة من المظاهر بحيث تمثل مجتمعة المعنى العام لصدق المحتوى ، وكل مظهر لهذا الصدق يكاد يكون شكل من أشكال صدق المحتوى أما هذه (المظاهر)الأنواع فهي :

الصدق الظاهري Face validity

وهو ابسط مظاهر صدق المحتوى ويسمى الصدق السطحي أو الخارجي بالنسبة للفقرات ويعني أن تكون الفقرات تقيس السمة (المحتوى) الذي تقيسه أو تنتمي للسمة التي تقيسها، ويتم توفيره من خلال حكم المختصين في مجال السمة المراد قياسها كأن يعطى اختبار للغة العربية لمعلم اللغة العربية ويقر بانتماء الفقرات للسمة أو الموضوع الذي تقيسه. وقد يطلق عليه اسم الصدق السطحي أو الصدق المنطقي كونه يشير إلى المظهر العام للاختبارمن حيث موضوعيته ووضوح فقراته وتعليماته ويتطلب هذا النوع من الصدق، البحث عما (يبدو) أن الاختبار يقيسه والتفحص المبدئي لمحتويات الاختبار وكذلك ومراجعة فقرات الاختبار والتحقق فيما إذا أنها تقيس تلك السمة

المعدة لقياسها أو ترتبط بالوظائف التي يحققها الاختبار. فإذا كان هناك ارتباط أو انتماء لمجال السمة المراد قياسها كان الاختبار صادقا بشكل ظاهري او بشكل أولي او مبدئي، وأما كيفية التعبير عن ذلك أو الكشف عنه فيتطلب التحليل المبدئي لفقرات الاختبار لمعرفة ما إذا كانت تتعلق بالسمة المقاسة وهذا أمر يرجع إلى ذاتية باني الاختبار (المعلم) وتقديره وهنا تكمن المحاذير التي تنتج عن ذاتية المعلم او المحكم ومدى اهتمامه بالكشف عن مؤشرات الصدق الظاهري.والذي يسمى أحيانا بصدق المحكمين الذين يحكموا مدى انتماء الاختبار إلى مجال السمة المراد قياسها ومن فوائد الانتباه إلى هذا المظهر من صدق المحتوى ضبط باني الاختبار على اثر التداخل المتوقع بين المواد او بين المواضيع في المادة الواحدة فمثلا قد يكون الاختبار معد لقياس المبتدا والخبر ويتداخل مع جمع المذكر السالم أو المثنى من حيث حركات الإعراب ، ويكتشف المحكم أن بعض الفقرات تقيس إعراب الجمع المذكر السالم أكثر مما تقيس المبتدأ والخبر ، وهذا بالطبع يتطلب الانتباه فقد يخل باني الاختبار بالصدق دون أن يشعر ولذلك يتم الحكم على الصدق الظاهري من خلال المختصين أو المحكمين حيث يطلق على هذا النوع من صدق المحتوى أحيانا بصدق المحكمين، ومن هنا فان الصدق الظاهري يركز على مضمون الفقرات المكونة للاختبار

- **الصدق المنطقي (العيني):**

يشير هذا النوع من صدق المحتوى إلى مدى تمثيل الفقرات المكونة للاختبار في صورته النهائية للعد اللانهائي من الفرات التي يمكن صياغتها لقياس المحتوى أو السمة المراد قياسها وبالتالي فان متطلب التمثيل أيضا ينطبق على الأهداف أو الكفايات الفرعية التي سيتم قياسها من السمة الأم. ولذلك يتوقف هذا النوع من الصدق على التحديد السليم للسمة المراد قياسها ، وهذا بالطبع يتحدد جزءا منه من خلال جدول المواصفات للاختبار من خلال مكونات كل موضوع فرعي وفي أي مجال من مجالات الأهداف والاهم من ذلك تحديد الأهمية النسبية لكل موضوع وكل مجال ولذلك من المهم أن ينتبه باني الاختبار أيضا هنا إلى انتقاء الطريقة المناسبة (المحك) الذي يتم الاعتماد عليه لتقرير الأهمية النسبية التي تحدد عدد الفقرات للموضوع المرتبطة به.

ومن المهم أن نعي أن انتهاك الصدق العيني يؤدي إلى الوقوع في الخطأ العيني الخاص بعينة الفقرات التي يتكون منها الاختبار في صورته النهائية على اعتبار أن الاختبار هو عينة من الفقرات لقياس عينة من السمات الفرعية المكونة للسمة الأم.وخلاصة القول أن الصدق العيني للاختبار يركز على عدد الفقرات المكونة للاختبار ومدى توزيعها على الموضوعات الفرعية ومجالات الأهداف

يتبين مما سبق أهمية صدق المحتوى والذي لا بد من التركيز فيه على عدد ومضمون فقرات الاختبار وهذا يتوفر من خلال الاهتمام بإعداد لائحة أو جدول المواصفات ، ومن المشكلات الرئيسية لصدق المحتوى انه لا يوجد مؤشرات إحصائية تشير إلى مدى توفر صدق المحتوى ، ومن هنا اعتمد بالدرجة الأولى على آراء المحكمين لكن يمكن زيادة الثقة به من خلال زيادة عدد المحكمين والاهتمام بمدى علاقتهم بالسمة التي يقيسها الاختبار وأخيرا فان صدق المحتوى للاختبار يخدم نفس الاختبار أي أن معيار الحكم هنا معيارا داخليا

ثانيا : الصدق المرتبط بالمحك (Criterion Related Validity)

يشير المعنى العام لهذا النوع من الصدق إلى الفائدة العملية أو الوظيفية للاختبار، بمعنى أن محك الحكم على صدق الاختبار هنا مدى تحقيقه للغرض الذي اعد من اجله ، أي أن مؤشر صدق الاختبار هنا هو مدى الارتباط بين نتائج المفحوصين على الاختبار قيد الاهتمام ونتائج نفس المفحوصين على اختبار آخر يقيس السمة أو سمة مرتبطة بها والمحك هو اختبار يقيس سمة أخرى لها علاقة بالسمة التي يقيسها الاختبار قيد الاهتمام. وبالنسبة لاختبار المحك يكون تم التحقق من صدقه لذلك نقارن بينه وبين الاختبار قيد عن طريق معامل الارتباط بينهما. ولذلك فالصدق المرتبط بمحك يعتمد على إيجاد معامل الارتباط بين الاختبار قيد الاهتمام واختبار آخر سبق إثبات صدقه. ومن هنا فانه كما أن لصدق المحتوى مظاهر فان للصدق المرتبط بمحك مظهرين أو نوعين ويتحدد كل نوع حسب الغرض من الاستخدام هما.وفي الصدق التنبؤي يتم الحديث عن نوعين من الصدق وفقا للغرض من استخدامه هما

- **الصدق التنبؤي Predictive Validity**

يدل هذا النوع من الصدق على مدى الصحة التي يمكن أن نتوقع بها خاصية أو قدرة معينة لدى الأفراد من خلال اختبار يفترض أن يقيس هذه الخاصية. يعتبر هذا النوع من الصدق مؤشرا لنتيجة معينة في المستقبل حيث يقوم على أساس المقارنة بين درجات الأفراد في الاختبار وبين درجاتهم على محك يدل على أدائهم في المستقبل ، ويعتبر معامل الارتباط بين درجاتهم على الاختبار ودرجاتهم على المحك هو معامل صدق الاختبار. وعليه فلتوفير مؤشر أو معامل الصدق يتم حساب الارتباط بين درجات المفحوصين على الاختبار وبين درجاتهم على المحك وبحجم معامل الارتباط يكون الصدق. وكمثل على ذلك فإذا خضع مجموعة من المتقدمين للقبول في التربية الرياضية في الجامعة وخضعوا لاختبار القدرات الجسمية والبدنية لكلية التربية الرياضية، وحصلوا على نتائج عالية ، وخلال الدراسة الجامعية وتحديدا في السنة الرابعة مثلا كانت معدلاتهم التراكمية عالية فان معامل الارتباط بين درجاتهم على اختبار القدرات ودرجاتهم في المساقات الجامعية الرياضية سيكون مؤشرا على الصدق التنبؤي ولو أمكننا إعداد اختبار لقياس قدرات الطلبة الذين انتهوا من دراستهم بالمرحلة الثانوية بهدف تحديد مدى النجاح الذي سيحققه هؤلاء الطلاب في دراستهم الجامعية ، وجاءت درجة الارتباط عالية بين تحصيل هؤلاء الطلاب في الجامعة وبين درجاتهم في اختبار القدرات بعد الثانوية، فإننا نستطيع القول أن هذا الاختبار له قدرة تنبؤية كبيرة ، وعندما نريد حساب مؤشر الصدق التنبؤي لا بد من حساب القيمة التنبؤية للاختباروالاعتماد على فكرة أن السمة المقاسة لها صفة الثبات النسبي في المواقف المستقبلية ويحتاج هذا النوع من الصدق (التنبؤي) إلى فترة بين تطبيق الاختبار وجمع المعلومات عن المحك في فترة تالية للاختبار

- **الصدق التلازمي Concurrent Validity**

يمثل الصدق التلازمي العلاقة بين الاختبار ومحك موضوعي تجمع البيانات عليه وقت أو قبل إجراء الاختبار . أي التعرف على مدى ارتباط الدرجة على الاختبار بمحكات الأداء الراهنة أو مركز الفرد حاليا. يستخدم عندما يتلازم تطبيق الاختبار وتطبيق المحك معا ويصبح الهدف هو معرفة عما إذا كان كل من الاختبارين يقيسان خصائص قائمة بالفعل في وقت واحد ، وذلك بهدف تقدير الحالة الراهنة وهو من انسب الأساليب ملائمة للاختبارات التشخيصية فإعداد اختبار لقياس السرعة لو ارتبط بدرجة أو تقدير المدرب أو المدرس لأفراد العينة ، فان معامل الارتباط العالي مؤشر صدق تلازمي .

فهو يعبر عن مدى الارتباط بين النتائج التي يتم الحصول عليها بواسطة أداة القياس التي أعدها الباحث (الاختبار) وبين النتائج التي يتم الحصول عليها بواسطة أداة أخرى ذات صدق مقبول ، ويتطلب الصدق التلازمي لأداة ما تطبيق تلك الأداة على المفحوصين وتطبيق الأداة الأخرى على نفس المفحوصين في نفس الوقت ، ثم إيجاد درجة الارتباط بين النتائج التي تم الحصول عليها بواسطة الأداتين ، ومعامل الارتباط الذي نحصل عليه في تلك الحالة يعبر عن الصدق التلازمي للأداة التي أعدها الباحث . ومن شروط المحك الجيد .

- أن يكون مرتبط بالسمة التي وضع الاختبار لقياسها .
- أن الاختبار كمحك غير كتحيز لفئة ما من المفحوصين ، أي يتيح العدالة للجميع.
- أن يتميز بالثبات المقبول.
- يكون يتميز المحك بالموضوعية.

ومن المحاذير التي يجب الانتباه إليها بالنسبة للصدق المرتبط بالمحك هو اعتماده على صدق الاختبار المرجعي (المحك) فإذا كان هذا الاختبار غير صادق أو مشكوك في صدقه فانه سيؤثر بذلك على الاختبار المراد معرفة صدقه ، ولذلك يمكن تصحيح مؤشر معامل الصدق في حال كان ثبات كل من الاختبار المحك أو المراد إيجاد صدقه

مشكوكا فيه أي يوجد فيهما أخطاء قياس وهو ما يسمى بتصحيح اثر الانخفاض لثبات أي من الاختبارين.حيث يؤثر انخفاض الثبات على دقة القرارات التي يمكن أن تتخذ بالاعتماد على الاختبار قيد الاهتمام أو المحك وتؤثر بالتالي على الصدق الذي يلعب أهمية كبيرة في نسبة الاختيار وتصنيفهم حسب المحك او الاختبار قيد الاهتمام حيث طور كل من تيلر ورسل(Taylor-Russell,1939) جداول خاصة لذلك تعتمد على كل من معاملي الثبات والصدق وعلى ما يسمى بمعدل الاختيار (Selection Rate) ومعدل الأساس (Base Rate)من حيث نسبة المختارين

ويمكن التمييز بين نوعي الصدق التنبؤي والتلازمي من خلال الفترة الزمنية بين الاختبار قيد الاهتمام والاختبار المحك ، والهدف من الاختبار فإذا كان الغرض تحديد الوضع الحالي كان الصدق اللازم هو الصدق التلازمي وإذا كان الغرض التنبؤ بنتيجة مستقبلية كان الصدق اللازم هو الصدق التنبؤي

ثالثا : صدق المفهوم (البناء) Construct Validity.

يقصد بصدق المفهوم مدى نجاح الاختبار في قياس مفهوم فرضي معين فمثلا الذكاء سمة غير مادية وغير محسوسة إنما يمكننا ملاحظة بعض المظاهر الخاصة بها ، والقلق له مظاهر مثل احمرار الوجه أو سرعة الكلام وعدم التركيز على سبيل المثال وقد نصمم مواقف اختبارية نقيس من خلالها مدى القلق وتحديد إن كان الطالب عالي القلق يؤدي مهمات صعبة أفضل من الطالب متدني القلق . وحتى نصل إلى ذلك لا بد من تطبيق مقياس للقلق على المفحوصين أو العينة المقصودة حتى يمكن تصنيفهم إلى مرتفعي القلق أو منخفضي القلق، وإذا كان الاختبار يقيس القلق ويتوصل إلى النتائج المفترضة والمعروفة يكون الاختبار صادقا من الناحية البنائية.إذ يعتمد باني الاختبار على نظريات خاصة بالسمة التي صمم الاختبار لقياسها ويصف على أن صدق البناء استمرارية لتطور السمة المقاسة، ولذلك إذا انسجمت نتائج الاختبار مع الافتراضات التي استند إليها مطور (باني) الاختبار فان ذلك مؤشر على صدق البناء، أما في حال عدم الانسجام فان ذلك يعني خطأ في التصميم التجريبي للتطبيق ، أو أن

هناك خطأ أو خلل في بنية النظرية المتعلقة بالسمة التي يقيسها الاختبار ، أو أن الاختبار لا يقيس السمة المصمم لقياسها.

ويتعلق مفهوم صدق البناء ببنية الاختبار ومكوناته ويسمى أحيانا بالصدق التكويني وهو من أكثر أنواع الصدق تعقيدا لأنه يعتمد على افتراضات نظرية يتم التحقق منها تجريبيا ويمكن الوصول إليه من خلال عدة أساليب مثل أسلوب المجموعات المتطرفة وأسلوب الاتساق الداخلي وأسلوب التحليل العاملي، ويعتبر التحليل العاملي من أفضل الأساليب للحصول على مؤشر لصدق البناء. حيث إلى دراسة الظواهر المعقدة ويعتمد فكرة الارتباط لاستخلاص العوامل التي تؤثر أو تكون الظاهرة أو السمة قيد الاهتمام من خلال تحليل معاملات الارتباط بين متغيرات الظاهرة بين المكونات، ويتم استخدام عدة طرق في ذلك منها طريقة المكونات الأساسية لهوتلنج (Hotleing) وهي الأكثر شيوعا بسبب دقة نتائجها، وهنا يتم بلورة عدة عوامل وتسمى بالعوامل المباشرة، حيث لا يمكن تفسيرها إلا بعد تدويرها (Rotation)، والتدوير عملية قائمة على أسس رياضية تهدف إلى تحقيق تركيب بسيط لمصفوفة معاملات الارتباط الداخلية بحيث ترفع قيمة التشبعات الكبيرة وتقلل من قيمة التشبعات الصغيرة، وللتدوير طريقتين الأولى طريقة التدوير المتعامد (الفاريماكس لكايزر) والتدوير المائل(الكوارتماكس)، وبعد استخراج العوامل يستخدم الباحث الاختبار إذا زاد تشبعه عن قيمة محددة وتتعلق بقيمة تسمى قيمة الجذر الكامن التي لا بد أن تكون اكبر من قيمة يحددها باني الاختبار حسب معرفته بطبيعة السمة التي يقيسها الاختبار ..الخ

كما لا بد من الإشارة إلى بعض المفاهيم المرتبطة بصدق البناء كونه يعتمد على مفهوم الارتباط كمؤشر عليه (الصدق) فمن المعروف أن ارتباط السمة مع نفسها يكون ارتباطا تاما ولذلك فانه كلما ارتفع الارتباط بين سمة وسمة أخرى ارتفعت قيمة معامل الارتباط وينطبق ذلك على نتائج الاختبارات والتي تقيس سمة واحدة أو عدة سمات ، فمعاملات الارتباط بين الاختبارات المختلفة التي تقيس سمة واحدة لا بد أن تكون عالية وهذا يعني ان هذه الاختبارات متلازمة أو متقاربة وهو ما يسمى

بالصدق التقاربي (Convergent Validity)، وعلى العكس فمعاملات الارتباط بين الاختبارات المختلفة والتي تقيس سمات مختلفة (مستقلة) لا بد أن تكون معاملات الارتباط بينها متدنية وهذا يعني أن هذه الاختبارات متباعدة من حيث العلاقة بينها وهو ما يسمى بالصدق التباعدي (Divergent Vhdity).

وللكشف عن صدق البناء للاختبار يرى جرونلاند (gronland,1966) ضرورة إتباع الخطوات التالية :

- التعريف بالإطار النظري للسمة المرتبطة بنتائج الاختبار.
- اشتقاق الفرضيات المرتبطة بنتائج الاختبار .
- اختبار صحة الفرضيات والتحقق من ذلك منطقياً أو تجريبياً .

ولا بد أن نذكر هنا إلى أن اختلاف معيار الحكم على أنواع الصدق وأهميتها ليس مدعاة للتقليل من أهمية كل نوع ، إنما يعتمد ذلك على طبيعة الاختبار ، فعندما ننظر إلى الاختبار كممثل لا بد من الاهتمام بصدق المحتوى أما عندما ننظر إلى الاختبار كمتنبئ فلا بد من الاهتمام بالصدق المرتبط بمحك، حيث أن الاختبار الممثل بالضرورة أن يكون متنبئا أما الاختبار المتنبئ فليس من الضروري أن يكون ممثلا حيث يتم الاهتمام بالارتباط بين نتائج الاختبار ونتائج اختبار آخر يقيس نفس السمة إذ تعتمد قدرته التنبؤية على قوة هذا الارتباط

صدق الاختبار محكي المرجع.

كنا قد أسلفنا الاختلاف بين الاختبار محكي المرجع والاختبار معياري المرجع من حيث الغرض والطبيعة والتفسير ولا تختلف أهمية مؤشر الصدق لكلا الاختبارين إنما قد تختلف من حيث المفهوم الخاص بهما ، حيث أن صدق الاختبارات محكية المرجع لم يحظ بالاهتمام الكافي كما هو الحال بالنسبة للثبات ، وبما أن الاختبار محكي المرجع يركز على وصف وتحديد السمة (النطاق السلوكي) المقاسة بشكل دقيق باعتبار أن كل فقرة تقيس كفاية أو هدفا تعليميا فان مفهوم الصدق للاختبار محكي المرجع لا يختلف كثيرا عنه للاختبار معياري المرجع إلا من حيث طبيعة الغرض الذي يصمم له

وان عدم الاختلاف الكبير لا يمنع من التطرق إلى أنواع الصدق الخاصة بالاختبار محكي المرجع وهنا يمكن أن نتحدث عن ثلاثة أنواع من الصدق على النحو التالي

الصدق الوصفي: Disprective Validity

يتميز الاختبار محكي المرجع بأنه يصف السمة المقاسة وصفا دقيقا بشكل أكثر منه بالنسبة للاختبار معياري المرجع ، وتبعا لذلك فإننا نستطيع تفسير النتائج بطريقة أدق من الاختبار معياري المرجع ، ويقابل الصدق الوصفي هنا صدق المحتوى وتتلخص وظيفته أو أهميته في وصف السمة ويمكن أن يكون أكثر عمومية من صدق المحتوى لذلك أطلق عليه الصدق الوصفي، ويمكن ان نوفر الصدق الوصفي بنفس الطريقة التي نوفر فيها صدق المحتوى أي من خلال المحكمين المختصين بالسمة موضوع الاختبار.

الصدق الوظيفي: Functional Validity

يهتم هذا الصدق بطبيعة الوظيفة التي صمم من اجلها الاختبار ، ويتعدى الوصف الدقيق للسمة المقاسة ، ويقابل هذا النوع الصدق المرتبط بالمحك أو الصدق التجريبي Empirical Validity ولكنه أكثر شمولية من حيث أن التنبؤ يتم من خلال الوصف الدقيق للسمة المقاسة والذي قد لا ينطبق على الدق المرتبط بالمحك كون التنبؤ يتم دون الاهتمام بالتحديد الدقيق للسمة المقاسة، أي أن التنبؤ يتم من خلال مواقف تختلف عن المواقف التي يقيسها الاختبار معياري المرجع وقد يقودنا هذا المعنى إلى التركيز على الصدق الوظيفي على حساب الصدق الوصفي

صدق الانتقاء للنطاق السلوكي للاختبار Domain Selection Validity

ويتعلق هذا النوع من الصدق بقواعد تحديد السمة المقاسة بمكوناتها المختلفة أي بطريقة تحليل الموضوعات الفرعية (الأبعاد) للسمة قيد القياس ، فقد نحلل السمة أو الموضوع الذي يقيسه الاختبار حسب المفاهيم أو الحقائق أو القيم أو القوانين وقد نحلله حسب مستويات الأهداف ، ونحصل في كل مرة على نمط من السمات الفرعية للسمة او الموضوع ونصمم اختبار حسب كل طريقة تحليل ونطبق الاختبارين على

مجموعة محددة ومعروفة في مستواها من حيث السمة المقاسة ونلاحظ النتائج في كلا التطبيقين ، وفي هذه الحالة التحليل الذي نعتقد انه يمثل السمة المقاسة أفضل تمثيل وبالتالي أفضل مجموعة من الفقرات وهذا ما يقصد بهذا النوع من الصدق.وبمعنى أكثر تحديد فان هذا النوع من الصدق يركز على التحديد الأفضل لمكونات السمة المقاسة.وأكثر ما يشبه صدق البناء للاختبار معياري المرجع.وقد اقترح بابام (Popham,1978) طرقا لاختيار التحديد الأدق للسمة موضوع القياس ، وتعتمد هذه الطرق على التجريب والوصول إلى الإتقان لمضمون الكفايات التي تقيسها الفقرات التي يمكن تصميمها لتصميم اختبار محكي المرجع

مؤشرات الصدق .

لقد أسلفنا بأن صدق الاختبار يشير إلى قدرته على قياس ما يرغب باني الاختبار أو المستخدم قياسه وبالتالي فان هذه القدرة لا بد أن يكون لها قيمة أو مؤشر يدلل على مداها أو توفرها ، ويعتبر معامل الارتباط من المؤشرات على الصدق والقيمة التي تعبر عن الصدق يمكن تسميتها بمعامل الصدق.وعلى الرغم من أن لمعامل الارتباط قيمة وإشارة تعبر عن اتجاهها إلا أننا هنا سنهتم بقيمة معامل الارتباط الموجبة كون المرغوب به هنا هو ارتباط طردي (موجب) وهذا يشير إلى أن قيمة معامل الصدق ستكون ضمن المدى (صفر – ١) حيث أن القدرة التنبؤية لا تعتمد على إشارة معامل الارتباط (اتجاهه) واما كيفية حساب معاملات الصدق فهي تعتمد بشكل رئيسي على معامل الارتباط ولذلك يصنف الصدق أحيانا في ضوء الإجراءات التي تستخدم لحساب مؤشراته مثل الصدق العاملي والصدق التنبؤي والصدق التقاربي والتباعدي (التميزي) على النحو التالي كل حسب الإجراءات الملائمة له.

صدق المحتوى.

يشير إلى بيان ما إذا كانت مفردات الاختبار تمثل المجال الذي وضع الاختبار لقياسه. وفي حساب صدق المحتوى يجب أن نضع في اعتبارنا درجة مناسبة نوع المفردات لقياس ما وضعت لقياسه ودرجة شمول عينة المفردات والطريقة التي تقاس

بها محتويات مجال الاختبار وللتأكد من توفر صدق المحتوى ينبغي الإجابة على الأسئلة التالية :

- هل يحتوى الاختبار على معلومات كافية لتغطية ما يفترض أنه يقيسه ؟
- هل أسئلة الاختبار مناسبة وهل الاختبار يقيس المجال المراد قياسه ؟
- ما مستوى الإتقان الذي يقاس به محتوى الاختبار ؟

وإذا تم الإجابة على تلك الأسئلة إجابات مرضية فان ذلك يعبر عن أن محتوى الاختبار جيد. ولا ينبغي الخلط بين صدق المحتوى والصدق الظاهري ، وأحيانا يستخدم الاتساق الداخلي للمفردات على انه صدق محتوى من خلال حساب معاملات الارتباط بين درجات الأفراد على كل فقرة ودرجاتهم في الاختبار ككل، وهذا يؤكد صدق المفردات وليس صدق المحتوى

الصدق الظاهري .

يمكن حساب الصدق الظاهري للاختبار عن طريق التحليل المبدئي لفقراته بواسطة عدد كبير من المحكمين لتحديد ما إذا كانت هذه الفقرات تتعلق بالجانب الذي تقيسه وهنا لا يمكن لباني الاختبار أن يقوم بعمل تكرارات استجابات مجموعة من المحكمين ويختار الفقرات التي اتفق عليها أكبر عدد من المحكمين ، ويشير الصدق الظاهري إلى السمة التي يظهر أن الاختبار يقيسها بشكل أولي (ظاهري) وبالرغم من أن هذه الطريقة ليست كافية للتأكد من صدق الاختبار ولكنها تفيد في طمأنة الباحث مبدئيا على دقة الاختبار الذي يستخدمه في مقياس ما صمم لقياسه.

الصدق المرتبط بمحك

ويتم حساب الصدق بهذه الطريقة بحساب مدى اتفاق درجات الأفراد على الاختبار الجديد (المراد حساب صدقه) ودرجاتهم على اختبار آخر سبق حساب صدقه وثباته ويقيس نفس جوانب السلوك التي يقيسها الاختبار الجديد ويعاب على تلك الطريقة أن معامل الصدق الناتج يعني ارتباط درجات الاختبار الجديد باختبار آخر قديم يقيس نفس ما يقيسه الاختبار الجديد من جوانب سلوكية ، وهذا يعني أن

الباحث لم يواجه مشكلة في قياس ما يريد قياسه من جوانب سلوكية بحيث يقوم بتصميم اختبار جديد خاصة وان تصميم الاختبارات وتقنينها من الأمور الشاقة للغاية.

الصدق التلازمي

ويمكن حساب معامل الصدق في هذه الطريقة بمعامل الارتباط بين درجات الأفراد على الاختبارات ودرجاتهم في الأداء العملي في جوانب السلوك التي يقيسها الاختبار . بشرط أن تكون درجات أداء الأفراد الفعلية قد تم جمعها في نفس الوقت الذي يطبق فيه الاختبار أو قبلها بفترة قليلة

الصدق التنبؤي

ويمكن حساب معامل الصدق في هذه الطريقة بمعامل الارتباط بين الدرجات على الاختبار ودرجات الأداء الفعلي للأفراد كما يقاس بطريقة أخرى بعد إجراء الاختبار ويقوم حساب المعامل هنا على حساب القيمة التنبوية للاختبار مثل حساب صدق اختبارات الاستعدادات الخاصة كالاستعداد الدراسي أو الرياضي أو الموسيقى

صدق البناء

ويمكن الاستدلال على هذا النوع من الصدق للاختبار من خلال حساب الارتباط بين درجات الأفراد على أبعاد الاختبار وبين مفهوم هذه الجوانب كما تحددها النظرية التي تبناها باني الاختبار في ن والتصميم ، أي انه في نهاية الأمر يرجع الفروق بين درجات الأفراد إلى اختلاف مستوياتهم في السمة التي تعالجها النظرية ويقيسها الاختبار ويستخدم أحيانا التحليل العاملي في التأكد من توفر صدق البناء للاختبار حيث تعتمد هذه الطريقة في حساب معامل صدق الاختبار على أسلوب تحليل إحصائي يسمى التحليل العاملى (Factor Analysis) حيث يهدف إلى تحديد مدى قياس مجموعة اختبارات لبعض العوامل المشتركة ونتيجة لذلك يسمى بالصدق العاملي نسبة للتحليل العاملي المستخدم في الكشف عنه، لكن من عيوب تلك الطريقة كثرة عدد معاملات الصدق العاملي للاختبار الواحد ، وذلك عندما يتشبع هذا

الاختبار بعوامل مختلفة ، وبهذا يكون الاختبار غير صادق لأنه يتشبع بعوامل أخرى ولا يقيس جوانب السلوك التي ينبغي أن يقيسها .

العوامل المؤثرة في الصدق:

- **عوامل خاصة بالاختبار:**
- طول الاختبار . حيث يزداد صدق الاختبار بزيادة فقراته .
- ثبات الاختبار : ويتأثر الصدق بقيمة الثبات ، لذلك فالنهاية العظمى للصدق لا تزيد عن الجذر التربيعي لمعامل ثبات الاختبار
- اللغة السهلة وبسيطة التي يجب أن تناسب ومستوى المفحوصين يزيد من صدقه.
- سهولة أو صعوبة الفقرات والتي تؤدي إلى حصول المفحوص على درجات لا يستحقها وهذا يضع تساؤلات تتعلق بصدق الاختبار،
- إشارة بعض الفقرات إلى الإيحاءات للإجابة مما يؤثر بصدق الاختبار
- **عوامل خاصة بتطبيق وتصحيح الاختبار**
- العوامل البيئية في إجابة المتعلم على فقرات الاختبار سلبا أو إيجابا وهذا قد يقلل أو يزيد أو يقلل من صدق الاختبار.
- مستوى مراجعة وطباعة وإخراج الاختبار يؤثر في صدق الاختبار
- تطبيق الاختبار مع مجموعات من الطلبة غير التي صمم ليتناسب مع مستواها وخصائصها .
- **عوامل خاصة بالمفحوص:**
- الاضطراب والقلق الزائد لدى المفحوص مما يؤثر في أدائه أي على على نتيجته.
- التخمين والغش من العوامل التي تؤثر على مستوى المتعلم في الاختبار
- تباين الفروق الفردية بين المفحوصين حيث أن ارتفاع وتدني مدى الفروق الفردية تؤثر على صدق الاختبار

دار جرير للنشر والتوزيع

العلاقة بين صدق الاختبار وثباته

تعتبر العلاقة بين صدق الاختبار وثباته، فكلاهما وجهان لشيء واحد هو مدى صلاحية ذلك الاختبار في أن يقيس ما وضع لقياسه وفي إعطائه نتائج متماثلة، إذ يفترض في الاختبار أن يكون صادقاً وثابتاً، ولذا يفترض أن تكون العلاقة بين كل منهما علاقة إرتباطية عالية ، وهناك مجموعة من العوامل تؤثر في صدق الاختبار وثباته منها تلك العوامل المتعلقة بالاختبار نفسه من حيث لغته، وإجراءات تطبيقه وتصحيحه، وصياغة فقراته، وسهولة تلك الفقرات أو صعوبتها، وطول الاختبار أو قصره، ومنها تلك العوامل المتعلقة بالمفحوص نفسه، ويقصد بها تلك العوامل وظروفه الصحية، ومنها تلك العوامل البيئية المتعلقة بشروط عملية تطبيق الاختبار مثل العوامل الفيزيائية كالإضاءة والتهوية ودرجة الحرارة (الروسان، ١٩٩٦) وبطبيعة الحال فان الصدق يحقق الثبات لكن العكس لا يتم ، ولذلك يمكن القول أن الاختبار الصادق يكون اختبارا ثابتا بالضرورة والعكس ليس صحيحا، كما ويمكن التعرف إلى طبيعة العلاقة بينهما من خلال الوعي بان الحد الأعلى لمعامل الصدق هو الجذر التربيعي لمعامل الثبات .كما أن الثبات يتأثر بتباين الخطأ اي باخطاء القياس وكذلك بنسبة التباين الحقيقي إلى التباين الكلي، بينما يتأثر معامل الصدق بالتباين الحقيقي، وهذا يعني أن معامل ثبات الاختبار يتمثل في نسبة التباين الحقيقي المنسوب وغير المنسوب للسمة التي يقيسها الاختبار ، في حين أن معامل الصدق يتمثل في نسبة التباين المنسوب فقط إلى السمة التي يقيسها الاختبار، وهذا يقودنا إلى التذكير بان زيادة طول الاختبار الأصل أن تزيد من طول الاختبار شريطة أن تحقق الزيادة في عدد الفقرات زيادة في تمثيل محتوى ذلك الاختبار، ونخلص إلى أن خاصيتي صدق وثبات الاختبار صفتين متلازمتين لذلك فالعلاقة بينهما تكاملية، لكن لا بد من أن نتذكر أن الصدق والثبات مفهومان لا يخضعان للكل أو العدم .

الباب الثالث

نظرية السمات الكامنة

الفصل الأول

الأطر النظرية

بقيت النظرية التقليدية في القياس (Clasrical Test Theory) سائدة حتى عقد الستينات من هذا القرن، حيث بدأت معالم النظرية الحديثة في القياس Item Respones Teory) ، إذ خرج إلى حيز الوجود نموذج يرتكز إلى افتراضات هذه النظرية وذلك على يد جورج راش (Georg Rasch,1960) حيث استندت هذا النموذج إلى علم الاحتمالات والنماذج المنبثقة عنه.

وإذا ما أمعنا النظر في المتغيرات قيد الاهتمام في المشهد التربوي وبشكل أكثر تحديدا في مجال القياس النفسي، فإننا نجد الكثير من التداخل بين هذه المتغيرات، والتي تعبر في واقع الحال عن سمات الأشياء أو الأفراد، كما أن كل متغير من هذه المتغيرات يتكون من مجموعة من المتغيرات أو السمات الفرعية أو الجزئية. والتي تتجمع(تتآزر معا) لتشكل متغيرا بحد ذاته، كما انه يجب أن يؤخذ بعين الاعتبار أن هذه المتغيرات تتصف بالتجريد أي أنها في الغالب غير محسوسة ويكتنفها بعض الغموض وهو ما أطلق عليه رواد النظرية الحديثة في القياس بصفة الكمون (Latent)، ومع ذلك فان الحديث عن متغير ما يحفزنا نحو تصور مكونات هذا المتغير، فإذا ما ورد متغير الذكاء مثلا (Intellegence) فسيتبادر إلى الذهن الذكاء العالي والمتوسط والمتدني رغم أننا لن نستطيع تحديد المكونات تحديدا دقيقا نظرا لكمون هذه المكونات، وكذلك بالنسبة لمتغير القدرة المدرسية (scholastic ability) ومكوناته حيث يمكن تذكر بعض المفاهيم أو القدرات ذات العلاقة مثل القدرة على استغلال الوقت أو الحصول على المعلومات ذات العلاقة بدقة وبطريقة سهلة أو القدرة على تعلم مواضيع أو مفاهيم معينة بسهولة وسرعة.

ومن المعلوم بأن الميدان التربوي وتحديدا العمل الأكاديمي يعج بالكثير من المتغيرات كالقدرة القرائية والرياضية .....الخ حيث يعتبر علماء النفس بان هذه المتغيرات عبارة عن خصائص أو سمات مجردة أو كامنة، أي أنها غير محسوسة بشكل مباشر بالرغم من إمكانية وصفها وتجزئتها إلى مكوناتها أو تفصيلها من الناحية النظرية، لكن قياسها وتحديد مقدارها في الواقع يعد أمرا فيه بعض الصعوبة، إذ أن قياس وتحديد مقدار الذكاء والتحصيل والدافعية لا ينطبق عليه قياس وتحديد مقدار الوزن والطول مثلا

إن الهدف الأساسي الذي يسعى إليه التربويون وعلماء القياس هو تحديد ما يمتلكه الفرد من سمة ما، ولذلك فان المواضيع التي يتناولها البحث العلمي تتمركز حول هذه المتغيرات، كالتحصيل المدرسي، القدرة القرائية، القدرة الرياضية.. ... الخ وعلى العموم فان مفهوم القدرة (Ability) في النظرية الحديثة للقياس يستخدم ليشير إلى مفهوم السمات الكامنة (Latent traits)، فإذا ما أردنا قياس أو تحديد ما يمتلكه فرد ما من سمة معينة فانه من الضروري توفر كل من وحدة القياس وأداة القياس كالأداة المستخدمة في قياس الطول وهي المتر بوحدة السنتمتر أو مشتقاته، وهذا يعني إمكانية تحديد أداة القياس وعدد الوحدات وبالتالي مقدار السمة، وإلا فانه من الصعب التعبير عن السمة بمجرد توفر عدد من الأرقام ومن هنا فان الفصول الأولى من هذا الكتاب ستعالج هذه القضايا من خلال تعريف ما يمكن تسميته مقياس القدرة (Ability Scale) والذي يفترض أنه ومهما كانت القدرة فانه يمكن قياسها (تحديدها) على مقياس يمتد ضمن المدى(∞، -∞) ويمثل نقطة (الصفر) منتصف هذا المقياس بمعنى أنه يمكن توفير وحدة قياس وكذلك خاصية الصفر المطلق والذي يمكننا من التعامل مع المتغيرات التي تقع على مستوى القياس الفتوي

أما الفكرة الأساسية لأداة القياس هذه فهي أنه يمكن قياس أو تحديد ما يمتلكه فرد ما من سمة ما، إضافة إلى إمكانية مقارنة قدرات الأفراد مع بعضهم البعض على متصل تتحدد قيمه النظرية بين (∞، -∞) أما المدى الفعلي له فيقع بين (٣- ،٣) وان القدرة قد تصل إلى الحد الأعلى أو الأدنى من مسطرة القياس



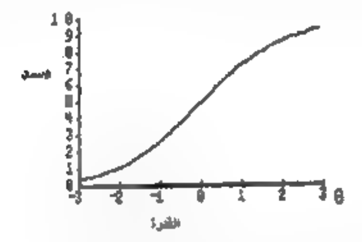
دار جرير للنشر والتوزيع

منحنى خصائص الفقرة

إن الوسيلة الأكثر ألفة أو استخداما لقياس القدرة فهي بناء أو تطوير اختبار مكون من مجموعة من الفقرات(Item) بحيث تقيس كل فقرة أو مجموعة من الفقرات جانبا معيناً من جوانب السمة المراد قياسها، ومن المهم وحسب النظرية الحديثة في القياس أن تتحرر إجابة المفحوص على فقرة معينة من إجابته على فقرة أخرى، وإذا كانت الإجابة صحيحة حينها يحصل المفحوص على العلامة(١) وإذا كانت الإجابة خاطئة حينها يحصل المفحوص على العلامة(صفر) وهذا يعني أن الفقرة لا بد أن تكون مدرجة تدريجا ثنائيا (صفر، ١) حيث تفترض نظرية القياس الحديثة بأن قدرة المفحوص تتحدد بعدد الفقرات التي يجيب عليها إجابة صحيحة، إذ أن كل مستوى قدرة يمكن صاحبه من الإجابة على الفقرات التي تتناسب ومستوى قدرته وهذا يعني أن المفحوص سيجيب عن الفقرات التي تتمتع بمستوى صعوبة لا يتعدى مستوى قدرته، وبالتالي فان العلامة الكلية للمفحوص تتحدد بعدد ونوعية الفقرات التي أجاب عنها إجابة صحيحة، وهو عكس ما تفترضه النظرية الكلاسيكية في القياس (Classical Test Theory) حيث تفترض أن العلامة الكلية للمفحوص هي مجموع الفقرات التي أجاب عنها المفحوص إجابة صحيحة بغض النظر عن مستوى صعوبتها. ومن هنا فانه من الصعب استخدام الفقرات المبنية على أساس النظرية الحديثة كما هو مألوف في النظرية الكلاسيكية لأنه من الصعب على المفحوص الإجابة على الفقرات بنمط معين لأن المفحوص الذي يجيب على فقرة ذات مستوى صعوبة ما لا بد أن يجيب على الفقرات ذات مستوى الصعوبة الأقل منها

ومن الافتراضات المنطقية والمعقولة للنظرية الحديثة في القياس أن مستوى القدرة الذي يتمتع به أي مفحوص يتحدد باحتمال أجابته عن أي فقرة ويزداد هذا الاحتمال بزيادة القدرة (ق) ويقل بانخفاضها، بمعنى أن العلاقة بين القدرة واحتمال الإجابة الصحيحة عن فقرة ما هي علاقة طردية وحيث أن مدى القدرة يقع ضمن المدى (٣- ٣) فان قيم احتمال الإجابة الصحيحة على الفقرة (ح س) تقع ضمن المدى (صفر - ١) وبما أن العلاقة بينهما علاقة طردية فانه يمكن التعبير عنها بصيغة رياضية لمنحنى هذه العلاقة باستخدام متصل القدرة ومقدار احتمال الإجابة الصحيحة، وبما أن العلاقة طر دية فان شكل منحنى

العلاقة يتخذ شكل الحرف (S) كما هو في الشكل (١٠) حيث يسمى بمنحنى خصائص الفقرة (Item Characteristic Curve) ويقصد بالخصائص هنا الصعوبة واحتمال الإجابة الصحيحة والتمييز إضافة إلى القدرة التي تحدد كل هذه المعلمات . ويعتبر مفهوم منحنى خصائص الفقرة المفهوم الأهم والأساسي لنظرية القياس الحديثة، حيث يعتمد عليه كل المهتمين بالقياس مصممين أو مستخدمين للاختبارات، ولذلك فهو الذي يتم التركيز عليه لدوره المهم في توضيح البناء النظري لهذه النظرية، ويتميز بخاصيتين أساسيتين حيث يستخدما لتوضيحه وهما:



شكل (١٠) منحنى خصائص لفقرة مقبولة منطقيا

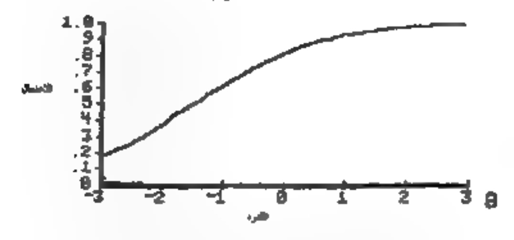
معلمة صعوبة الفقرة (Difficulty) :

وهي معلمة يتم التعبير عنها عادة بمقدار احتمال الإجابة الصحيحة على الفقرة مقترنا بمستوى القدرة الذي يمتلكه المفحوص وسنرمز له بالرمز (ص).

معلمة تمييز الفقرات (Discrimination) :

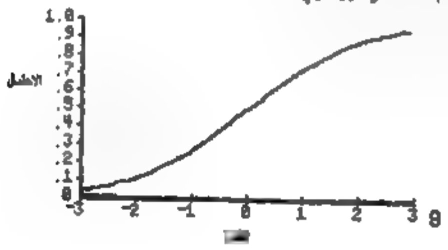
وهي معلمة يتم التعبير عنها عادة بقدرة الفقرة على التمييز بين المفحوصين الذين تجتاز قدرتهم الإجابة على الفقرة إجابة صحيحة والمفحوصين الذين لا تمكنهم قدرتهم من الإجابة على الفقرة إجابة صحيحة وسنرمز له بالرمز (ت)، وهذا يتطلب ملاحظة منحنى خصائص الفقرة من القسم الأوسط ويوضح الشكل رقم (١١ أ، ب، ج) مفهوم الصعوبة





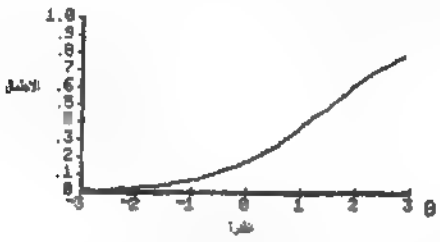
شكل(١١أ) منحنى الخصائص لفقرة سهلة وتمييزها مثالي

حيث يشير كل شكل منها إلى فقرة مستقلة تتمتع بمعلمة ذات صعوبة مختلفة ومتقاربة في معلمة التمييز، فالفقرة التي يمثلها المنحنى (١١ ب) فقرة سهلة والسبب في ذلك هو زيادة احتمال الإجابة عليها إجابة صحيحة والذي تمثله القيم على محور الصادات وذلك بالنسبة للمفحوصين ذوي القدرة المتدنية والعالية معا، أما الفقرة التي يمثلها المنحنى (٢ب) فهي فقرة متوسطة الصعوبة حيث أن احتمال الإجابة الصحيحة عليها يقترب من القيمة (٥ ٠) وذلك بالنسبة للمفحوصين ذوي القدرة المتوسطة وهو منخفض بالنسبة للمفحوصين ذوي القدرة المتدنية



شكل(١١ب) منحنى الخصائص لفقرة متوسطة الصعوبة وتمييزها مقبول

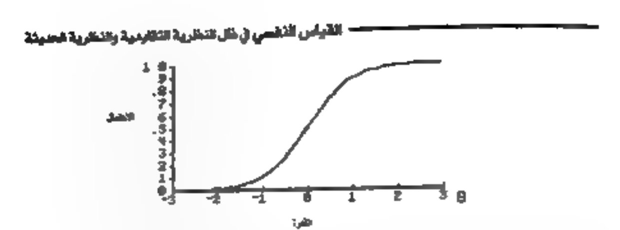
أما الفقرة التي يمثلها المنحنى (١١ج) فهي فقرة صعبة بسبب انخفاض احتمال الإجابة الصحيحة عليها عند غالبية المفحوصين ومن ذوي مستويات القدرة العالية لاحظ الزيادة الملحوظة لقيمة احتمال الإجابة الصحيحة على الفقرة مع زيادة مستوى القدرة حيث يبلغ احتمال الإجابة الصحيحة عليها (٠.٨) حتى عند أعلى مستوى قدرة (ق =٣).



شكل(١١ج) منحنى الخصائص لفقرة متوسطة الصعوبة وتمييزها مقبول

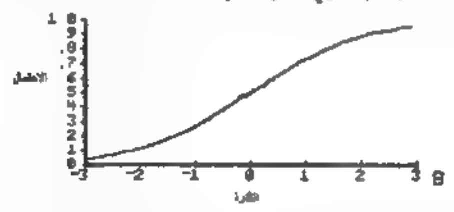
أما بالنسبة لخاصية التمييز فتتضح من خلال الشكل رقم (١٢، أ، ب، ج د) والذي يشير الى فقرات لها نفس مستوى الصعوبة ومختلفة من حيث قدرتها التمييزية، فالفقرة التي يمثلها المنحنى رقم (١٢ أ) لها قدرة تمييزية عالية حيث الميل الواضح للمنحنى عند مستوى قدرة (صفر) ويتغير المنحنى (يرتفع) بشكل ملحوظ كلما زاد مستوى القدرة ويستدل على القدرة التمييزية العالية للفقرة من خلال ميل المنحنى .





شكل(١٢ أ) منحنى الخصائص لفقرة ذات تمييز عالي وصعوبة أعلى من المتوسطة

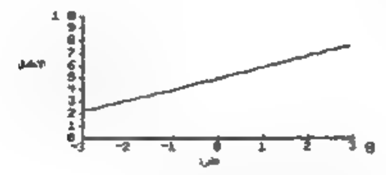
أما الفقرة التي يمثلها المنحنى رقم (١٢ب) فان قدرتها التميزية متوسطة حيث أن ميل المنحنى أقل منه للفقرة ذات المنحنى (أ)، إضافة إلى أن زيادة احتمال الإجابة الصحيحة عليها يزداد بشكل منسجم مع الزيادة في مستوى القدرة.وكما نلاحظ فان احتمال الإجابة الصحيحة عليها يقترب من (صفر) بالنسبة لذوي القدرة المتدنية ويقترب من (١.٠) بالنسبة لذوي القدرة العالية .



شكل(١٢ب) منحنى الخصائص لفقرة ذات تميز متوسط وصعوبة متوسطة

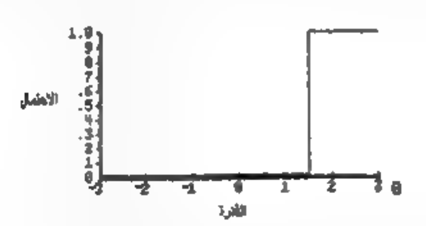
أما الفقرة التي يمثلها المنحنى رقم (١٢ج) فان قدرتها التمييزية متدنية حيث انخفاض احتمال الإجابة الصحيحة عليها لمستوى القدرة المتوسط ويزداد قليلا لمستوى

القدرة العالية وبالتالي فهي فقرة غير مميزة بين المفحوصين من مستويات القدرة المختلفة.



شكل(١٢ج) منحنى الخصائص لفقرة ذات تمييز متوسط وصعوبة متوسطة

أما الفقرة التي يمثلها الشكل ١٢د) فان قدرتها التمييزية عالية (تمييز تام) حيث يلاحظ الارتفاع الملحوظ والعامودي لمنحنى الخصائص عند مستوى القدرة (ق = ١.٥)، كما أنها فقرة صعبة لجميع المفحوصين من ذوي مستويات القدرة الأقل من (١.٥) وهي تميز بين مجموعتين من المفحوصين فوق وتحت (ق= ١.٥)



شكل(١٢ د) منحنى الخصائص لفقرة ذات تمييز تام وصعوبة عالية.

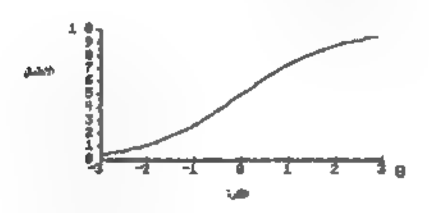
وحسب نظرية الاستجابة على الفقرة (Item Response Theory) وتفصيلا لمنحنى خصائص الفقرة فان الصعوبة والتمييز تقسم إلى أربعة مستويات حسب الجدول رقم (٤) المبين أدناه .

جدول (٤) مستويات معلمتي الصعوبة والتمييز حسب نظرية الاستجابة على الفقرة.

الخاصية	المستويات				
الصعوبة	سهلة جدا	سهلة	متوسطة	صعبة	صعبة جدا
التمييز	غير مميزة	منخفضة	متوسطة	عالية	تامة

وسيتم إيراد مجموعة من الفقرات التي توضح الخاصيتين الأساسيتين لكل فقرة من حيث الصعوبة والتمييز كما هو في الشكل (١-٤ - أ، ب، ج).

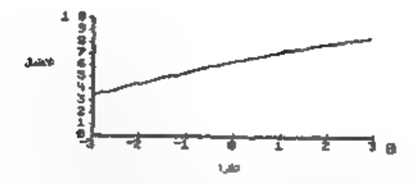
ففي الشكل (١٣ أ) فقرة متوسطة الصعوبة(ص) والتمييز (ت) وبالنسبة للصعوبة فان احتمال الإجابة على الفقرة إجابة صحيحة يزداد بزيادة مستوى القدرة وأما بالنسبة للتمييز فان منحنى خصائص الفقرة يبدأ بالتصاعد (يزيد الميل) بشكل ملحوظ عند مستوى القدرة (ق = صفر) أي أن هذه الفقرة تميز بين مجموعتي المفحوصين فوق وتحت مستوى (ق= صفر)



شكل (١٣) منحنى الخصائص لفقرة متوسطة الصعوبة والقدرة التمييزية

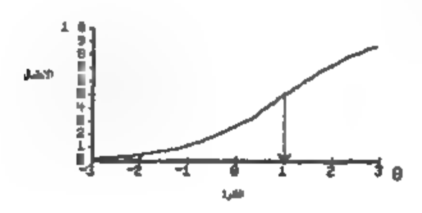
وفي الشكل (١٣ب) يتضح منحنى الخصائص لفقرة سهلة لان احتمال الإجابة الصحيحة عليها مرتفع نوعا ما عند مستويات القدرة المتدنية ويزداد بزيادة مستوى القدرة أما التمييز فهو متدني حيث أن التغير في منحنى خصائص الفقرة (الارتفاع) غير متمايز ويكاد يكون خطا مستقيما لغاية مستوى القدرة المتوسطة (ق) = (صفر) وهذا مؤشر على أنها لا تميز بين المفحوصين.

دار جرير للنشر والتوزيع



شكل (١٣ ب) منحنى الخصائص لفقرة سهلة ومتدنية القدرة التمييزية

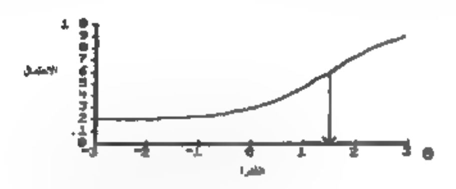
وفي الشكل (١٣ج) يتضح منحنى الخصائص لفقرة أعلى من المتوسط من حيث الصعوبة لان احتمال الإجابة الصحيحة عليها يزداد بشكل متناغم مع الزيادة في مستوى القدرة حيث أن احتمال الإجابة عليها(٠.٥٠) عند مستوى القدرة (ق =١.٠٠)، لكنه لا يبلغ الحد الأعلى أو يقترب منه حتى مع زيادة مستوى القدرة أما بالنسبة للتمييز فهو متوسط حيث أن منحنى خصائص الفقرة ذو شكل مفلطح وليس حاد أي أن ميل المنحنى متزايد لكن بشكل قليل ويصل إلى الثبات النسبي عند مستويات القدرة العالية



شكل (١٣ ج) منحنى الخصائص أعلى من متوسطة الصعوبة ومتوسطة القدرة التمييزية

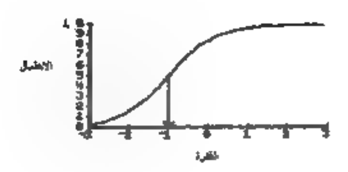


وفي الشكل (١٣د) يتضح منحنى الخصائص لفقرة صعبة لان احتمال الإجابة الصحيحة عليها يبقى منخفض حتى مع زيادة مستوى القدرة، أما التمييز فهو متوسط لان ميل المنحنى يبدأ بالزيادة عند مستويات القدرة العالية ق= (٢.٠) . بمعنى آن المفحوصين تحت وفوق مستوى القدرة (ق = صفر) لديهم القدرة على إجابة الفقرة إجابة صحيحة، والتمييز يبدأ عند مستوى (ق= صفر) فهي لا تميز بين المفحوصين من مستويات القدرة ق = (- ٣) ـ ق = (١).



شكل (١٣ د) منحنى الخصائص لفقرة صعبة ومتدنية القدرة التمييزية

وفي الشكل (١٣هـ) يتضح منحنى الخصائص لفقرة سهلة نوعا ما لزيادة احتمال الإجابة الصحيحة عليها حتى عند مستوى القدرة (ق= صفر)، أما التمييز فهو تام وذلك لان احتمال الإجابة الصحيحة على الفقرة يزداد بشكل ملحوظ من (-٢.٠) إلى (صفر) أي أن ميل المنحنى يرتفع بشكل ملحوظ عند القدرة (-٠ ١) لكن الفقرة لا تميز بين المفحوصين من هم أقل من مستوى القدرة (-٢.٠) ومن هم أعلى من مستوى القدرة ق=(٠ ١) لأن المنحنى خطي تقريبا بعد مستوى القدرة (ق=١)، بمعنى أنها تميز بين المفحوصين من هم أعلى أو أقل من مستوى قدرة (- ١.٠)



شكل (١٣ هـ) منحنى الخصائص لفقرة صعبة وتامة القدرة التمييزية

لاحظ أن احتمال الإجابة الصحيحة (ح=٥٠ ٠) يقابل مستوى قدرة (ق= - ١.٠) وهو مقدار صعوبة الفقرة، وميل المنحنى عند نقطة تقاطع خط أفقي باتجاه خط الاحتمال (المحور الصادي مع المتحنى والخط العامودي بين المنحنى وخط متصل القدرة (المحور السيني يمثل تمييز الفقرة).

خلاصة

- يكون شكل منحنى خصائص الفقرة مسطحا (مفلطحا) إذا كانت القدرة التمييزية للفقرة متوسطة
- يأخذ منحنى خصائص الفقرة شكل حرف(S) أي يكون حميقا من منطقة الوسط، إذا كانت القدرة التمييزية للفقرة أعلى من مستوى التميز المتوسط.
- إذا كانت صعوبة الفقرة أقل من المتوسط فان احتمال الإجابة عليها إجابة صحيحة يكون أكبر من (٠.٥) .
- إذا كانت صعوبة الفقرة أعلى من المتوسط فان احتمال الإجابة عليها إجابة صحيحة يكون أقل من (٠.٥) .
- إذا تم تمثيل الفقرة بحسب الصعوبة والتمييز فلا يغيب عن البال أن كل منهما مستقل عن الآخر.
- إذا كانت الفقرة لا تميز بين المفحوصين فان الصعوبة ولجميع المفحوصين تقع على خط أفقي وذلك عند الاحتمال (ح س = ٠.٥) وهذا يكون بسبب عدم تعريف أو تحديد صعوبة الفقرة التي لا تميز بين المفحوصين أي الإجابة عليها تتم بالتخمين.
- يمكن تحديد صعوبة الفقرة من النقطة التي تتقاطع مع منحنى خصائص الفقرة إذا ما رسمنا خطا أفقيا من نقطة الاحتمال (٥ ٠)، حيث يمثل هذا احتمال الإجابة الصحيحة للفقرة السهلة بالنسبة للمفحوصين من ذوي مستوى القدرة المتدنية، كما يمثل احتمال الإجابة الصحيحة للفقرة الصعبة بالنسبة للمفحوصين من ذوي مستوى القدرة العالية.

الفصل الثاني

نماذج منحنى خصائص الفقرة

Item Characteristic Curve Models

في الفصل الأول من هذا الباب تم الحديث عن منحنى خصائص الفقرة بشكل نظري أو من الناحية المفاهيمية، أما في هذا الفصل فسيتم عرض رؤية واضحة حول أنواع أو نماذج منحنى خصائص الفقرة وهذا ما تفرضه أهمية الوعي بنظرية السمات الكامنة (نظرية القياس الحديثة).وسيتم الحديث عن ثلاثة نماذج رياضية لمنحنى خصائص الفقرة باستخدام مفهوم اللوغريتمات (Logarithms) إذ يتضمن كل نموذج منها معادلة رياضية تبين العلاقة بين احتمال الإجابة الصحيحة على الفقرة (ح س)والقدرة (ق)، كما يتضمن كل نموذج واحدا أو أكثر من معلمات الفقرة (الصعوبة، التمييز، واحتمال الإجابة بالتخمين) وتتضمن كل معلمة منها قيمة رياضية تعبر عن المنحنى العملي لخصائص الفقرة، ويذكر أن كل نموذج من هذه النماذج قد لعب دورا في تطوير نظرية القياس الحديثة، ومن المهم أن نكون على دراية بأن كل نموذج من هذه النماذج يقدم حجما ما من المعلومات حول خصائص الفقرة، كما وأن لكل نموذج منها معادلة لحساب احتمال الإجابة الصحيحة على الفقرة عند مستويات مختلفة من القدرة، إذ يمكننا ذلك من رسم شكل منحنى خصائص الفقرة، ومن هنا فإن هذا الفصل يهدف إلى تزويد القارئ بالصورة الواضحة حول مساهمة كل قيمة من قيم إحصائيات(معلمات) الفقرات في إخراج منحنى الخصائص وذلك حسب كل نموذج بمعادلته وإجراءاته الحسابية الخاصة به. ومن الجدير ذكره أن أسماء هذه النماذج اقترنت بعدد المعلمات الخاصة بالفقرة والتي يتم الحصول عليها من إجراءات كل نموذج.

الاقتران اللوغريتمي The Logistic Function

إن النموذج الرياضي (المعياري) المعتمد بالنسبة للنظرية الحديثة في القياس والذي يصف منحنى خصائص الفقرة هو شكل تراكمي لمعادلة لوغريتمية ينبثق عنها مجموعة من المنحنيات التي تصف خصائص الفقرة، ومنها ما تم عرضه من أشكال في الفصل الأول، وقد كان أول توظيف لمفهوم اللوغريتمات والنماذج المنبثقة عنه عام (١٨٤٤)، ومن ثم بدا بالانتشار على نطاق واسع وفي علم الأحياء تحديدا، وذلك من اجل نمذجة نمو النباتات والحيوانات منذ الولادة وحتى الشيخوخة، وقد استخدم مفهوم اللوغريتمات لأول مرة كنموذج لمنحنى خصائص الفقرة في أواخر الخمسينيات وذلك لبساطته وسهولته حيث استخدم النموذج ثنائي المعلمة، أما هذه النماذج فهي على النحو التالي

أولا، النموذج اللوغريتمي احادي المعلمة One Parameter Model أو (Rasch Model)

وهو النموذج الأبسط من بين نماذج منحنى خصائص الفقرة ويعرف بنموذج جورج راش (Georg Rasch) عالم الرياضيات الدنمركي حيث كان أول من قام بنشره ولذلك اقترن هذا النموذج باسمه منذ الستينات من القرن العشرين، إذ استخدم نظرية الاحتمالات في تحليل البيانات، وكان هذا الإتجاه مغايراً لما كان مألوفا آنذاك، حيث أن النموذج الناتج لمنحنى خصائص الفقرة هو نموذجا لوغريتميا، وسوف نتناول هذا النموذج بشيء من التفصيل في الفصل الثامن من هذا الكتاب، أما في هذا الفصل فسينصب الاهتمام على نماذج منحنى خصائص الفقرة وحسب هذا النموذج سيتم تثبيت قيمة معلمة التميز (ت) على افتراض أنها متساوية لجميع الفقرات وقيمتها العددية تساوي (١) أما المعلمة التي سوف تأخذ قيما مختلفة فهي معلمة الصعوبة (ص) ولذلك سمي هذا النموذج بالنموذج أحادي المعلمة.أما الصيغة العامة لهذا النموذج فهي كما في المعادلة (٤٥)



ح (ق) = ١/ ١+ ث $^{-\text{د(ق-ص)}}$ ....(٤٥)

وبما أن قيمة معلمة التمييز قيمة ثابتة لجميع الفقرات فان المعادلة (٤٥) ستصبح على النحو.

ح (ق) = ١/ ١+ ث $^{-١(ق-ص)}$ ....(٤٥)

حيث

ح (ق) :احتمال الإجابة الصحيحة على الفقرة

ق: مستوى القدرة

ث: ثابت ومقداره ٢.٧١٨

ص: معلمة الصعوبة

وبما أن هذا النموذج يفترض أن معلمة التمييز ثابتة لجميع الفقرات ومساوية للقيمة (١) فان الحد الأيمن للمعادلة سيكون ($^{-١(ق-ص)}$) وهذا يبقي على الحد على النحو $^{-(ق-ص)}$ وهذا يعني أنه وبناء على افتراض هذا النموذج بان قيمة ث مساوية للقيمة (١) فان(ث) لم تظهر في المعادلة على أساس أن الحد مضروبا بالقيمة ١ ولذلك ظهر الحد الأيمن وهو - (ق- ص) كما هو في المعادلة.وسيتم عرض مثال يوضح بنية وإجراءات هذا النموذج.

مثال رقم (١)[*] :

في اختبار للرياضيات أعد لقياس تحصيل الطلبة في وحدة الاحتمالات، تم سحب إحدى الفقرات وكانت معلمة الصعوبة لها (ص = (١.٠)، فما هو احتمال الإجابة الصحيحة عليها من قبل مفحوص من مستوى القدرة (ق = (-٣.٠) ؟

* أخذ من كتاب. (the Basica ofitem Response + heory, Baker, 2001)

الحل:

ح (ق) = ١/ ١+ ث $^{-١(ق-ص)}$

ل = ت(ق- ص) . = ١(-٣ - ١) = ١(-٤) = -٤.

لكن الصيغة العامة تريد إيجاد قيمة - ل ومن هنا فان سل = -(-٤) = ٤ بالتطبيق في المعادلة (٢ -١) .

ح (ق) = ١/ ١+ ث $^{-١٠(ق-ص)}$

ح (ق) = ١/ ١+ ٢.٧١٨ $^{-١(-٣-١)}$

ح (ق) = ١/ ١+ ٢.٧١٨ $^{(٤)}$

ح (ق) = ١/ ١+ ٥٤.٥٩٨

ح (ق) = ١/ ٥٥.٥٩٨

ح (ق) = ٠.٠١٧٩

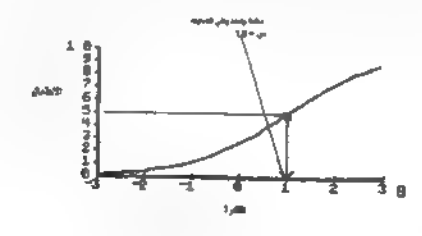
ح (ق)= ٠.٠٢

وإذا ما أردنا معرفة احتمال الإجابة الصحيحة على هذه الفقرات عند مستويات القدرة الأخرى وضمن المدى المعروف (-٣ - ٣) فإنها ستكون حسب الجدول رقم (٥). حيث تكرر الإجراءات الحسابية السابقة عند كل مستوى مع الأخذ بعين الاعتبار تغير قيمة (ق)

جدول(٥) :احتمال الإجابة الصحيحة لفقرة صعوبتها (١) وتمييزها(١) عند مستويات قدرة مختلفة وحسب النموذج أحادي المعلمة

القدرة	الصعوبة	التمييز	الانحراف (ل)	ث - الانحراف (- ل)	١ + ث - الانحراف (- ل)	الاحتمال ح(θ)
٣-	١	١	٤-	٥٤ ٥٩٨	٥٥ ٥٩٨	٠.٠٢
٢-	١	١	٣-	٢٠ ٠٨٦	٢١.٠٨٦	٠ ٠٥
١-	١	١	٢-	٧.٣٨٩	٨.٣٨٩	٠ ١٢
صفر	١	١	١-	٢ ٧١٨	٣ ٧١٨	٠ ٢٧
١	١	١	صفر	١.٠	٢.٠	٠.٥٠
٢	١	١	١	٠.٣٦٨	١.٣٦٨	٠ ٧٣
٣	١	٠ ٥	٢	٠ ١٣٥	١ ١٣٥	٠.٨٨

وإذا ما أمعنا النظر بقيم احتمال الإجابة الصحيحة على الفقرة وربطناها مع مستويات القدرة فإننا نلاحظ انه كلما زاد مستوى القدرة فان احتمال الإجابة الصحيحة يزيد وهذا امر منطقي ولذلك فان إجابة المفحوص على فقرات من مستويات صعوبة مختلفة يأخذ نمط معين من الاستجابات وإذا حاولنا رسم منحنى خصائص الفقرة في المثال السابق فسيكون كما هو في الشكل (١٤)



شكل (١٤) منحنى الخصائص للفقرة الواردة في المثال رقم (١)

ثانيا :النموذج اللوغريتمي ثنائي المعلمة Tow Parameter Model

وهو النموذج الذي يعتمد معلمتين هما معلمة الصعوبة (Difficulty) ومعلمة التمييز (Discrimination) وكان هو أول نموذج يستخدم في التطبيقات العملية لنظرية السمات الكامنة، ويتلخص بالمعادلة رقم (٤٦)

ح (ق) = ١/ ١+ ث $^{-ل}$ ....(٤٦).

حيث ل = ت(ق– ص) .

ويعبر عن (ل) بالصيغة ل = ت(ق– ص) ومنها – ل = – (ت(ق – ص)

وعليه فان المعادلة (٢–٢) تصبح على الصورة التالية

ح (ق) = ١/ ١+ ث $^{-ت(ق-ص)}$

ح (ق) = ١/ ١+ ث $^{-ت(ق-ص)}$ .. . .

حيث:

ق : مستوى القدرة.

ح (ق) : احتمال الإجابة على الفقرة إجابة صحيحة.

ث : قيمة ثابتة وقيمتها ٢.٧١٨ .

ص : صعوبة الفقرة.

ت : التمييز.

ل : الانحراف اللوغريتمي.

إن المعلمة الخاصة بالصعوبة والتي رمزنا لها بالرمز (ص) يمكن تعريفها بأنها النقطة التي تقع على متصل القدرة وذلك عندما يكون احتمال الإجابة على الفقرة إجابة صحيحة مساويا للقيمة (٠.٥) ويمكن أن تقع ضمن المدى النظري(– ∞ ـــ ∞)

وضمن المدى العملي (-٣ ــ ٣) وبالنظر إلى منحنى خصائص الفقرة الذي يأخذ شكل الحرف (S) فان ميله يزداد بزيادة مستوى القدرة، كذلك فان القدرة التمييزية لا تمثل الميل العام لمنحنى خصائص الفقرة كما تم تعريفه في الفصل الأول، أما التعريف التقليدي له فهو مدى التناسب بين ميل منحنى خصائص الفقرة عندما تساوى كل من القدرة والصعوبة، أما الميل الواقعي (الحقيقي) للمنحنى في حالة التساوي هذه فهي مساوية للمقدار (ت/ ٤) أي التمييز مقسوما على العدد(٤) لذلك فالمدى النظري لهذه المعلمة فهو (- ∞ ــ ∞) أما المدى العمل فهو (-٨ ٢ – ٢.٨). وستقوم بعرض مثال توضيحي لتوضيح النموذج اللوغريتمي ثنائي المعلمة لمنحنى خصائص الفقرة على النحو التالي.

مثال رقم (٢)

في اختبار لقياس تحصيل الطلبة في مادة العلوم كانت معلمة الصعوبة لاحدى فقراته كانت (ص=١) ومعلمة التمييز لنفس الفقرة (ت=٠.٥) احسب احتمال الإجابة الصحيحة على هذه الفقرة عند مستوى قدرة (ق = -٣)

الحل:

من خلال الصيغة العامة للنموذج فان هناك مجاهيل لا بد من إيجاد قيمها أولا ومن ثم إيجاد احتمال الإجابة على الفقرة إجابة صحيحة عند مستوى القدرة المحدد في المثال

ح (ق) = ١/ ١+ ث ^-ل^

و حيث أن (ل) = - ت(ق-ص) ......الانحراف اللوغريتمي.

إذن (ل) = - ٠.٥ (-٣ – ١)

= - ٠.٥(-٤.٠) = ٢.٠-

لاحظ أن المطلوب (- ل) ونحن حسبنا (ل) إذن (- ل) = -(-٢) = (٢)

والآن نطبق الصيغة العامة للنموذج لحساب الاحتمال

ح(ق) = ١/ ١+ ث ^-ل^

المقدار ث$^{-ل}$ = (٢٧١٨)^{٢}

= (٧.٣٨٩)

اذن بتطبيق معادلة النموذج العام ينتج ان

ح$^{(ق)}$ = ١/ ١+ ث$^{-ل}$

ح$^{(ق)}$ = ١/ ١ + ٧.٣٨٩

ح$^{(ق)}$ = ١/ ٨.٣٨٩

ح$^{(ق)}$ = ٠.١٢

وهذا يعني أن احتمال أن يجيب المفحوصين من ذوي مستوى القدرة (-٣) إذا كانت معلمة صعوبة (١) وتمييزها (٠.٥) ح$^{(ق)}$ (٠.١٢) وبناء على ذلك يمكن حساب احتمال الإجابة الصحيحة على هذه الفقرة عند مستويات القدرة ضمن المدى (-٣ ــ ٣) على النحو المبين في الجدول رقم (٦) وذلك بإعادة الإجراءات السابقة.

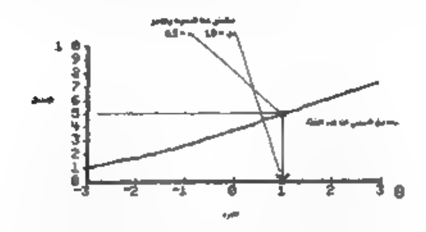
جدول(٦): احتمال الإجابة الصحيحة لفقرة صعوبتها (١) وتمييزها(٠.٥) عند مستويات قدرة مختلفة

القدرة	الصعوبة	التمييز	الانحراف (ل)	ث - الانحراف (-ل)	١+ث - الانحراف (ث-ل)	الاحتمال ح(ق)
-٣	١	٠.٥	-٢.٠٠	٧.٣٨٩	٨.٣٨٩	٠.١٢
-٢	١	٠.٥	-١.٥	٤.٤٨٠	٥.٤٨٠	٠.١٨
-١	١	٠.٥	-١.٠	٢.٧١٨	٣.٧١٨	٠.٢٧
صفر	١	٠.٥	٠.٥	١.٦٤٩	٢.٦٤٩	٠.٣٨
١	١	٠.٥	صفر	١.٠	٢.٠	٠.٥١
٢	١	٠.٥	-٠.٥	٠.٦٠٧	١.٦٠٧	٠.٦٢
٣	١	٠.٥	-١.٠	٠.٣٦٨	١.٣٦٨	٠.٧٣

يلاحظ أن احتمال الإجابة الصحيحة على الفقرة يزداد بزيادة مستوى القدرة ولذلك فان البيانات في الجدول أعلاه هي بمثابة معلومات عن الفقرة عند كل مستوى



قدرة ومن المعلومات يمكن أن ترسم منحنى خصائص الفقرة حيث سيكون على النحو المبين في الشكل (١٥).



شكل (١٥) منحنى خصائص الفقرة الواردة في المثال (٢)

نلاحظ أن هنا احتمال الإجابة الصحيحة على الفقرة يزداد بشكل متناغم مع الزيادة في مستوى القدرة، لكن التمييز غير واضح لأن ميل المنحنى منخفض ويظهر ذلك من خلال شكل المنحنى الذي هو عبارة عن شكل مستقيم نسبيا مع الارتفاع التدريجي إذ أن تميز هذه الفقرة متساوي عند جميع مستويات القدرة.

النموذج ثلاثي المعلمات : Three Parameter MOdel

جاء هذا النموذج ليغطي بعض الثغرات التي قد تظهر عند استخدام النموذجين. الأحادي والثنائي المعلمات، حيث يلجأ المفحوص إلى التخمين (Guessing) للإجابة على بعض الفقرات في حال عدم توفر القدرة الحقيقية للإجابة، وهذا يعني أن احتمال الإجابة الصحيحة سيتضمن جزءا سببه التخمين وليس القدرة الحقيقية على الإجابة، إن هذه الظاهرة لم يتم اخذها بالحسبان في النموذجين أحادي وثنائي المعلمة، وهذا ما جاء به بيرنبيوم (Birnbum,1968) إذ عدّل على النموذجين بحيث تضمن تعديله معلمة خاصة بأثر التخمين في احتمال الإجابة الصحيحة على

الفقرة، ونتيجة لهذا التعديل فقد النموذجين بعض الخصائص الرياضية وبهذا جاء النموذج الذي سمي بالنموذج ثلاثي المعلمات وتتمثل الصيغة العامة الخاصة به بالمعادلة (٤٧).

$$ح(ق) = تخ + (١-تخ)\ ١/\ ١+ ث^{-(ق-ص)}$$

حيث

ح (ق) :احتمال الإجابة الصحيحة على الفقرة

ق: مستوى القدرة.

ث: ثابت ومقداره ٢.٧١٨.

ص: معلمة الصعوبة.

تخ : احتمال الإجابة بالتخمين.

(١ – تخ) : متممة احتمال الإجابة بالتخمين.

إن المعلمة الجديدة في هذا النموذج هي (تخ) وتعرف على أنها احتمال الإجابة الصحيحة على الفقرة باستخدام التخمين فقط، أي بغياب القدرة الحقيقية على الإجابة، ومن المهم أن نأخذ بعين الاعتبار إن قيمة (تخ) لا تختلف باختلاف القدرة، وهذا يعني أن المفحوصين من مستويات القدرة المتدنية والعالية لديهم نفس الاحتمال للإجابة على الفقرة إجابة صحيحة بالتخمين لأن إجاباتهم تتم بغياب القدرة الحقيقية، وتقع هذه القيمة ضمن مدى نظري يقع بين (صفر ــ ١) أما القيمة العملية لهذا الاحتمال فهي دون القيمة (٠.٣٥) ولذلك فان هذه القيمة (٠.٣٥) هي القيمة العملية المسموح بها. ومن هنا فان دخول معلمة التخمين عدّل على مفهوم معلمة الصعوبة (ص)، ففي النموذجين الأحادي والثنائي المعلمة تم تعريف معلمة الصعوبة على أنها النقطة التي تقع على خط مستوى القدرة عندما يكون احتمال الإجابة

الصحيحة مساويا للقيمة (٠.٥) على افتراض أن قيمة معلمة التخمين تساوي (تخ = صفر)، أما في هذا النموذج فقد تم تحديد منحنى خصائص الفقرة بسبب أن التخمين أكبر من القيمة (صفر) أي أن (تخ > صفر) وعندما يكون احتمال الإجابة على الفقرة إجابة صحيحة مساويا للقيمة (٠.٥) أو مساويا لمعلمة صعوبة (٠.٥) فان الصيغة العامة لمعادلة النموذج ثلاثي المعلمات سيكون حسب المعادلة (٤٨) .

ح (ق) = تخ + (١-تخ) ٠.٥ ....(٤٨)
= ١ + تخ/٢

حيث أن ح (ق) هي احتمال الإجابة الصحيحة على الفقرة مع وجود احتمال التخمين ومن خلال المعادلة فان احتمال الإجابة الصحيحة سيقع بين قيمة التخمين (تخ) والقيمة (١) بدلا من المدى (صفر ــ ١) كما كان في النموذجين أحادي وثنائي المعلمة، ولذلك فان التخمين يعرف على انه الحد الأدنى من احتمال الإجابة الصحيحة على الفقرة، وقد نتج عن هذا التعريف أن معلمة الصعوبة (ص) تقع على خط القدرة عندما يكون احتمال الإجابة الصحيحة على الفقرة بين قيمة الإجابة بالتخمين والقيمة (١) .

أما بالنسبة لمعلمة التميز فقد احتفظت بتعريفها على أنها الانحدار النسبي لمنحنى خصائص الفقرة بالنسبة للقدرة التي تساوي معلمة الصعوبة أي أن (ق = ص)، وفي هذا النموذج (ثلاثي المعلمات) فان ميل منحنى خصائص الفقرة يحسب من المعادلة (٤٩)

م = ت(١ - تخ)/ ٤ ....(٤٩)

وذلك عندما تكون قدرة المفحوصين مكافئة لصعوبة الفقرة أي أن (ق = ص). وسنقدم مثالا يوضح بنية النموذج ثلاثي المعلمات وكذلك إجراءات إيجاد احتمال الإجابة الصحيحة على الفقرة إجابة صحيحة في ضوء قيمة معلمة التخمين كما في المثال التالي.

دار جرير للنشر والتوزيع

مثال رقم (٣)

خضع مجموعة من المفحوصين لاختبار يقيس تحصيلهم في مادة القياس والتقويم، وتم سحب فقرة تقيس معرفة الطلبة بالنموذج ثلاثي المعلمات وكانت هذه الفقرة تتمتع بقدرة تمييزية قيمتها (ت = ١.٣) ومستوى صعوبة (ص = ١.٥)، كما كان احتمال الإجابة على الفقرة بالتخمين (تخ= (٠.٢) فما هو احتمال الإجابة الصحيحة على هذه الفقرة عندما يكون مستوى القدرة (ق = -٣)؟

الحل :

المعطيات

(ص = ١.٥)، (ت = ١.٣)، (تخ = ٠.٢)

الخطوة الأولى ستكون إيجاد القيمة اللوغريتمية (ل)

القيمة اللوغريتمية(ل) = ت(ق- ص).

= ١.٣(- ٣ - ١.٥) .

= ١.٣(- ٤.٥) .

= ⁻ ٥.٨٥.

وبما أننا نريد – ل كما هي في الصيغة العامة فان – ل = ٥.٨٥

الخطوة الثانية ستكون إيجاد القيمة

(ث ^-ل^).

=(٢.٧١٨ ^٥.٨٥^). = ٣٤٧.٢٣٤.

الآن نجد القيمة (١ + ث ^-ل^).

= ١ + ٣٤٧.٢٣٤ = ٣٤٨.٢٣٤ .

والآن نطبق الصيغة العامة لإيجاد احتمال الإجابة الصحيحة على الفقرة حسب المعطيات.

$$\text{ح (ق)} = \text{١} / \text{١} + \text{ث}^{\text{ حط(ق- ص)}}$$

$$= \text{١} / \text{٣٤٨.٢٣٤}$$

$$= \text{٠.٠٠٢٩}$$

من الملاحظ انه وحتى هذه الخطوة فان لعمليات الحسابية لحساب الاحتمال هي نفسها بالنسبة للنموذجين الأحادي والثنائي المعلمة وذلك عندما (ص = ١.٥) و(ت = ١ ٣)، لكن الجديد هو أن هناك طرفا ثالثا دخل في معادلة النموذج ثلاثي المعلمة، ولإيجاد احتمال الإجابة الصحيحة على الفقرة حسب النموذج الثلاثي ستكون حسب المعادلة (٢ـ٣).

$$\text{ح (ق)} = \text{خ} + (\text{١}-\text{خ})\ \text{١} / \text{١} + \text{ث}^{\text{ حط(ق- ص)}}$$

مع ملاحظة أن قيمة الطرف الثالث من المعادلة.(١/ ١+ ث $^{\text{حط(ق- ص)}}$) = (٠.٠٠٢٩)

ومن هنا فان ح (ق) = ٠.٢ + (١- ٠.٢) (٠.٠٠٢٩) .

= (٠.٢) + (٠.٨) (٠.٠٠٢٩) .

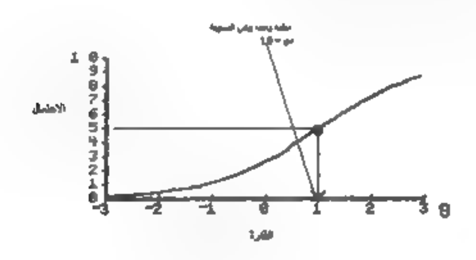
= (٠.٢) (٠.٠٢٣). = (٠.٢٠٢٣).

وهذا يعني أن احتمال الإجابة الصحيحة على الفقرة عند مستوى قدرة (ق = ٣) يساوي (٠.٢٠٢٣) ويبين الجدول رقم (٧) قيم احتمال الإجابة الصحيحة على الفقرة موضوع المثال حسب النموذج ثلاثي المعلمات وذلك عند مستويات قدرة مختلفة وبنفس الطريقة السابقة.

جدول (٧): قيم احتمال الإجابة الصحيحة على الفقرة موضوع المثال عند مستويات قدرة مختلفة وحسب النموذج ثلاثي المعلمات.

الفقرة	الانحراف (ل)	- الانحراف (- ب)	-الانحراف (ب-ل) +١	الاحتمال ح(θ)
٣-	٥.٨٥ -	٣٤٧.٢٣٤	٣٤٨ ٢٣٤	٠.٢٠
٢-	٤ ٥٥ -	٩٤ ٦٣٢	٩٥ ٦٣٢	٠ ٢١
١-	٣ ٢٥-	٢٥ ٧٩	٢٦ ٧٩	٠.٢٣
صفر	١ ٩٥-	٧.٠٢٩	٨.٠٢٩	٠ ٣٠
١	٠.٦٥-	١.٩١٦	٢ ٩١٦	٠.٤٧
٢	٠ ٦٥	٠ ٥٢٢	١ ٥٢٢	٠ ٧٣
٣	١ ٩٥	٠ ١٤٢	١ ١٤٢	٠ ٩٠

وإذا أردنا رسم منحنى خصائص الفقرة حسب المعلومات الواردة في الجدول فسيكون كما هو في الشكل (١٦) إذ أن لدينا ثلاث معلمات يمكن تحديدها على هذا المنحنى، لاحظ الشكل

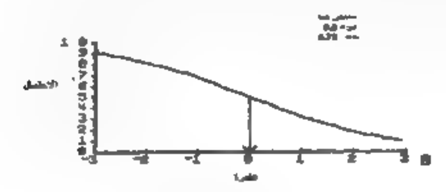


شكل (١٦) شكل منحنى الخصائص للفقرة الواردة في المثال (٣)

دار جرير للنشر والتوزيع

ظاهرة التمييز السالب Negative Discrimination :

إن التميز المرغوب به أو الطبيعي هو التميز الموجب ويحدث ذلك عندما يزداد احتمال الإجابة الصحيحة على الفقرة بزيادة مستوى القدرة، لكن هناك بعض الفقرات التي تظهر قدرة تمييزية سالبة، وفي هذه الحالة فان احتمال الإجابة الصحيحة على الفقرة يقل مع زيادة مستوى القدرة وذلك كما هو مبين في الشكل (١٧) .

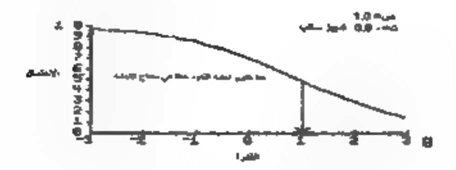


شكل (١٧) منحنى خصائص لفقرة ذات تميز سالب .

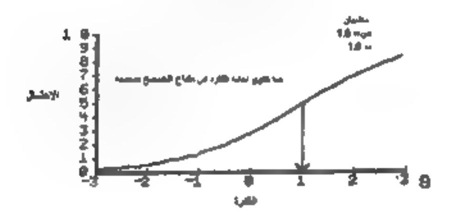
وتبدي الفقرة قدرة تمييزية سالبة في الحالتين التاليتين.

- إذا كانت الفقرة من نوع الاختيار من بديلين والإجابة الصحيحة دائما قيمة موجبة أي أنها تأخذ القيمة (١) .
- إذا كانت الفقرة من نوع الاختيار من متعدد وحينها فان هناك احتمالين هما.
  - الفقرة ضعيفة من حيث الصياغة.
  - الفقرة غامضة ومعلوماتها ناقصة.

والشكل (١٨أ، ب) يتضمن منحنى الخصائص لفقرتين في حالة الإجابة الصحيحة والإجابة الخاطئة.



شكل (١٨أ) منحنى الخصائص لفقرة تمييزها سالب في حال تكون الإجابة خاطئة في مفتاح الإجابة.



شكل (١٨ب) منحنى الخصائص لفقرة تمييزها موجب في حال تكون الإجابة صحيحة في مفتاح الإجابة.

ومن الشكل (١٨أ، ب) نلاحظ أن الفقرتين لهما نفس قيمة معلمة الصعوبة وكذلك معلمة التميز لكن مع اختلاف الإشارة بالنسبة لمعلمة التمييز فالفقرة رقم (١٨أ) لا تميز بين المفحوصين لأن احتمال الإجابة الصحيحة عليها يقل بزيادة مستوى القدرة للمفحوصين بينما تميز الفقرة رقم (١٨ب) بين المفحوصين حيث يزداد احتمال الإجابة الصحيحة عليها بزيادة مستوى القدرة للمفحوصين ويبين الجدول رقم (٨) التقديرات العددية لمستويات معلمة التميز والتقدير اللفظي المقابل لها

دار جرير للنشر والتوزيع

جدول (٨) : مستويات وأوصاف معلمة التميز .

الرقم	قيمة معلمة التميز	مستوى معلمة التميز
١	صفر	لا تميز
٢	٠.٠١ ـــ ٠.٣٤	ضعيف جدا
٣	٠.٣٥ ـــ ٠.٦٤	ضعيف
٤	٠.٦٥ ـــ ١.٣٤	متوسط
٥	١.٣٥ ـــ ١.٦٩	جيد
٦	أكبر من ١٧٠	جيد جدا
٧	+ ∞	ممتاز

إن القيم الواردة في الجدول رقم (٨) تصلح لتفسير معلمة التميز طبقا للنموذج اللوغريتمي (النظرية الحديثة) أما إذا أردنا تفسير معامل التمييز وفقا للنظرية الكلاسيكية (Classical Test Theory) فان ذلك يمكن أن يتم وذلك بعد قسمة كل قيمة من القيم الواردة في الجدول على القيمة (١.٧) (Lord ,1980) . فعلى سبيل المثال قيمة التميز المتوسطة حسب النظرية الحديثة والتي تقع ضمن المدى (٠.٦٥ ـــ ١.٣٤) على المنحنى اللوغريتمي فإنها تقع ضمن المدى (٠.٣٨ ـــ ٧٩ ٠) حسب مفهوم النظرية الكلاسيكية لمعامل التمييز، أما بالنسبة لإيجاد قيم مكافئة لمعلمة الصعوبة في حسب النظرية الحديثة أي على النموذج اللوغريتمي والنظرية الكلاسيكية فانه يؤدي إلى بعض المعضلات، حيث أن مفهوم الصعوبة ومتممته (السهولة) حسب النظرية الكلاسيكية فان الفقرة تقسم المفحوصين إلى مجموعتين تكون الفقرة لإحداهما صعبة وللأخرى سهلة، أما حسب النظرية الحديثة ونماذجها اللوغريتمية فان مفهوم الصعوبة عبارة عن نقطة تقع على متصل القدرة عندما يكون احتمال الإجابة الصحيحة على الفقرة (ص = صفر) بالنسبة للنموذجين الأحادي والثنائي المعلمة و (ص = ١ + تخ/ ٢) بالنسبة للنموذج ثلاثي المعلمة، ومن هنا فان مفهوم السهولة والصعوبة اللذين وردا في الفصل الأول فيحملان نفس المعنى وذلك عند نقطة منتصف متصل القدرة أي عندما (ص = صفر) . ومن هنا فان الطريقة المناسبة لتفسير

القيم الرقمية لمعلمة الصعوبة هي ملاحظة منحنى خصائص الفقرة على متصل (سمة القدرة .

أما بالنسبة لمعلمة التمييز فان هناك معنى آخر يضاف إلى التفسير السابق وهو أن ميل منحنى خصائص الفقرة عند أعلى مستوى قدرة يمثل أو يوازي الصعوبة لتلك الفقرة حيث تكون الفقرة أكثر تمييزا بين المفحوصين عندما تتساوى كل من قيمة معلمة الصعوبة ومستوى القدرة والسبب في ذلك أن قيمة ميل منحنى خصائص الفقرة (الذي يمثل معلمة التمييز) أعلى ما يمكن، ولذلك يمكن القول أنه وعند هذا المستوى من القدرة تكون قيمة معلمة الصعوبة (ص = - ١) بالنسبة للمفحوصين متدنيين مستوى القدرة وكذلك (ص = ١) بالنسبة للمفحوصين مرتفعي مستوى القدرة، هذا بالنسبة للنموذجين أحادي وثنائي المعلمة، أما بالنسبة للنموذج ثلاثي المعلمة فان قيمة معلمة التخمين (تخ) يتم تفسيرها على اعتبار أنها احتمال، فمثلا عندما تكون قيمة التخمين (تخ = ١٢ ٠) فان احتمال حصول المفحوصين على الإجابة الصحيحة على الفقرة من خلال التخمين هو (١٢ ٠) أي أن احتمال الإجابة على الفقرة إجابة صحيحة ولكن بفعل التخمين وليس نتيجة للقدرة الحقيقية للمفحوصين عند مستوى قدرة معين.

خلاصة

- إن ميل منحنى خصائص الفقرة حسب النموذج أحادي المعلمة يكون ثابتا، إلا عند النقاط التي يتغير فيها المنحنى
- في النموذج ثنائي المعلمة فان قيمة معلمة التمييز (ت) يجب أن تكون اكبر من (١.٧) وذلك عند النقطة التي تسبق منطقة التناقص للمنحنى.
- في النموذجين أحادي وثنائي المعلمة فان أعلى قيمة موجبة لمعلمة الصعوبة تكون عند منطقة التناقص (المنطقة المتدنية من المنحنى) من متصل القدرة أي كلما اتجهنا نحو القيمة (صفر)، أما في النموذج ثلاثي المعلمة فان أعلى قيمة موجبة للصعوبة تتحقق كلما اتجهنا نحو قيمة (احتمال) الإجابة بالتخمين.
- في النموذج ثلاثي المعلمة فان قيمة معلمة التخمين تقل كلما كانت قيمة معلمة الصعوبة متدنية (اقل من صفر) وكذلك كلما كانت قيمة معلمة التمييز متدنية(أقل من ١)، وإذا ما توزعت مستويات القدرة على متصل القدرة فان اقل قيمة لتناقص منحنى خصائص الفقرة تحدث عند قيمة معلمة التخمين .
- في النماذج اللوغريتمية الثلاثة فان المنحنيات التي تكون فيها قيمة معلمات التمييز سالبة ستكون انعكاسا للمنحنيات التي تكون فيها قيم معلمة التمييز موجبة.
- عندما تكون قيمة معلمة الصعوبة (ص = – ٣) فان النصف الذي سيظهر من منحنى خصائص الفقرة هو النصف العلوي، وفي حال كانت قيمة معلمة الصعوبة (ص = ٣) فان النصف الذي سيظهر من المنحنى هو النصف السفلي.
- إن اقل قيمة لميل منحنى خصائص الفقرة يحدث عند المناطق العميقة للمنحنى وذلك عندما تتساوى كل من الصعوبة والقدرة
- تعرف الصعوبة بأنها النقطة التي تقع على متصل القدرة عندما يكون احتمال

الإجابة الصحيحة على الفقرة (ح = صفر)، وذلك حسب النموذجين أحادي وثنائي المعلمات بينما تعرف الصعوبة حسب النموذج ثلاثي المعلمات على أنها النقطة التي تقع على متصل القدرة وذلك عندما تقع قيمة احتمال الإجابة الصحيحة على الفقرة بين احتمال الإجابة على الفقرة إجابة صحيحة بالتخمين والقيمة (١) أي ان (تخ) تقع ضمن المدى (تخ ـــ ١)، وتتساوى هاتين القيمتين أي قيمة (ح) والقيمة (١) عندما تكون قيمة احتمال الإجابة بالتخمين مساوية للصفر أي (تخ = صفر).

الفصل الثالث

تقدير معلمات الفقرة

إن الغرض الرئيسي لتحليل اختبار ما طبقا لإجراءات النظرية الحديثة في القياس هو تقدير معلمات فقرات الاختبار، إذ أن القيم الحقيقية لهذه المعلمات لا تكون معروفة قبل تطبيق الاختبار، حيث انه وبمجرد تقدير قيم هذه المعلمات فمن الممكن أن تزودنا هذه التقديرات بالمعلومات عن خصائص فقرات الاختبار، ويتم هذا التقدير بافتراض معرفة علامات القدرة للمفحوصين، حيث انه وفي حقيقة الأمر فان هذه العلامات تكون غير معروفة لكن من السهل وصف كيفية تحقيق تقديرات معلمات الفقرات لهذا الافتراض.

إن الوضع الطبيعي لأي اختبار إن يجيب عدد (عينة) من المفحوصين (ن) على عينة من الفقرات (ف) حيث تتوزع قيم علامات القدرة ضمن مدى مستويات القدرة التي تقع عادة على متصل القدرة، وبهذا فان عينة المفحوصين تنقسم إلى عدد من المجموعات الجزئية (ج) حيث سيكون لكل مجموعة من هذه المجموعات مستوى قدرة معينة(ق) حيث ستميز هذه المجموعات بمستوى القدرة الذي تمتلكه، كما انه ولكل مجموعة جزئية فان هناك عدد (مجموعة) من المفحوصين هم اللذين سيجيبون إجابة صحيحة على الفقرة بشكل صحيح (ج ي)، وتبعا لذلك فانه ولأي مستوى قدرة فان هناك احتمال ملاحظ للإجابة الصحيحة سيتم حسابه من المعادلة (٥٠)

ح(ع ه) = ع ه / ع ه ....(٥٠)

حيث:

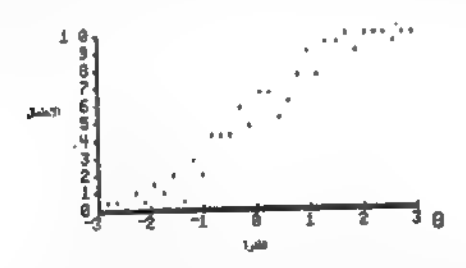
$$ح(ق_{ع}) = ج_{ق} / ج_{ه} \quad ....(٣ - ١)$$

ح(ق ع) الاحتمال الملاحظ للإجابة الصحيحة ولمستوى قدرة معينة (ق) ومن مجموعة (ج)

ج ق: عدد المفحوصين الذين أجابوا إجابة صحيحة على الفقرة من المجموعة ن.

ج ه: العدد الكلي لمجموعة المفحوصين في المجموعة ج.

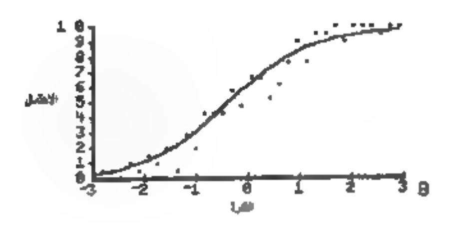
ومن الجدير بالذكر أن هذا الاحتمال يمكننا من تقدير احتمال الإجابة الصحيحة على الفقرة، كما ويمكننا تحديد عدد الذين أجابوا إجابة صحيحة على الفقرة (ج ق) وكذلك يمكن حساب الاحتمال الملاحظ ح(ق ع) ولكل مجموعة عند أي مستوى من مستويات القدرة، وإذا ما تم رسم شكل الاحتمال الملاحظ للإجابة الصحيحة لدى كل مجموعة فان النتيجة ستكون شكلا يكمل منحنى خصائص الفقرة كما في الشكل (١٩).



شكل (١٩) شكل الاحتمال الملاحظ للإجابة الصحيحة على فقرة.

والآن فان المهمة الرئيسية هي إيجاد منحنى خصائص الفقرة الذي يتطابق مع الاحتمالات الملاحظة للإجابات الصحيحة على فقرات الاختبار. وللقيام بذلك لا بد

من تحديد النموذج اللوغريتمي الذي نريد أن نطابق منحنى خصائص الفقرة به، وفي هذا السياق يمكننا استخدام أي نموذج من النماذج الثلاثة التي اشرنا إليها في الفصل الثاني وتعتمد الإجراءات هنا على أعلى نسبة ترجيح للتقدير (Maximum Liklihood ratio)، وفي ضوء ذلك فان القيم الأولية لمعلمات الفقرة هي معلمة الصعوبة (ص = صفر) ومعلمة التمييز (ت = ١.٠) حيث يتم تحديدها سلفا ومنهما يتم تقدير قيمة احتمال الإجابة الصحيحة على الفقرة، وذلك عند كل مستوى من مستويات القدرة المختلفة وذلك من خلال المعادلة الخاصة بالنموذج المراد استخدامه، حيث يجب إيجاد جميع القيم الملاحظة والقيم المحسوبة لجميع مستويات القدرة، بعد ذلك تتم إجراءات تعديلية (فلترة) على تقديرات معالم الفقرات، حيث أن الفقرة التي يجيب عليها المفحوصين إجابة صحيحة تبدي انسجاما أكثر مع منحنى خصائص الفقرة من خلال التقديرات التي تمت والاحتمالات الملاحظة للإجابة الصحيحة وتستمر هذه التعديلات حتى نصل إلى الحالة التي تبدي أكثر درجة من الاتفاق (المطابقة) وإذا وصلنا إلى هذه المرحلة (المطابقة التامة) تكون القيم التي وصلنا إليها هي التقديرات المعتمدة لكل من معلمات الصعوبة والتميز للفقرات وعندها يمكن تعويض هذه القيم في معادلة منحنى خصائص الفقرة وحسب النموذج المستخدم لإيجاد أو حساب احتمال الإجابة الصحيحة على الفقرة عند أي مجموعة وفي كل مستوى قدرة، وبذلك يمكننا رسم منحنى خصائص الفقرة إذ يُفترض أن يكون الأكثر تطابقا مع النموذج المستخدم، والشكل (٢٠) يبين منحنى خصائص فقرة متطابق مع القيم الاحتمالية الملاحظة للإجابة الصحيحة على الفقرة المبينة في الشكل (١٩)، والتي بلغت قيم تقديرات معلمة الصعوبة (ص= - ٠.٣٩) ولمعلمة التميز (ت = ١.٢٧)



شكل (٢٠) منحنى الخصائص لفقرة تطابق فيه الاحتمالات الملاحظة للإجابة الصحيحة على فقرة .

ومن الاعتبارات المهمة والتي يجب أخذها بالاعتبار في النظرية الحديثة في القياس أنه سواء بحثنا في مطابقة نموذج منحنى خصائص الفقرة مع بيانات الاستجابة على الفقرة أو مطابقة القيم الاحتمالية الملاحظة للإجابة الصحيحة مع تلك القيم المحسوبة فان الإحصائي المستخدم لقياس ذلك التطابق هو الإحصائي(χ^2) والملحق الخاص به حيث أن الصيغة العامة كما هي في المعادلة (٥١)

$$\text{كا}^2 = \sum_{\text{ج}=1} \text{ن}_{\text{ج}}(\text{ق}) \cdot \frac{\left[\text{ح}(\text{ق}_{\text{ج}}) - \text{ح}(\text{ق}_{\text{ج}})\right]^2}{\text{ح}(\text{ق}_{\text{ج}})\left[1 - \text{ح}(\text{ق}_{\text{ج}})\right]} \quad \ldots (٥١)$$

حيث



ج : عدد مجموعات القدرة.

ق م: مستوى القدرة للمجموعة م.

ج ه : عدد المفحوصين اللذين يحملون نفس مستوى القدرة.

ح(ق م) : الاحتمال الملاحظ للإجابة الصحيحة للمجموعة ه.

ح(ق م) الاحتمال المحسوب للإجابة الصحيحة لنفس المجموعة، ويحسب من النموذج اللوغريتمي المستخدم باستخدام تقدير المعلمات.

وبناء على ذلك يتم مقارنة القيم الملاحظة (الجدولية) بالقيمة المحسوبة فإذا كانت المحسوبة أقل من الملاحظة فهذا يعني أن منحنى خصائص الفقرة الذي سيتم رسمه من خلال بيانات الاستجابة الخاصة بالفقرة غير مطابق هذه للنموذج، ويحدث ذلك بسبب.

أولا: خطأ في نموذج منحنى خصائص الفقرة المستخدم.

ثانيا : القيم الاحتمالية الملاحظة للإجابة الصحيحة تنتشر (تتفرق) ولا تحقق تطابقا مع النموذج المستخدم.

وهذا لا يحدث في معظم الاختبارات، إذ أن بعض الفقرات تكون قيمة (كا²) المحسوبة عالية (تحقق التطابق) وعلى الغالب إذا أبدت غالبية الفقرات عدم التطابق مع منحنى خصائص الفقرة فان الخطأ يكمن في اختيار النموذج اللوغريتمي. وغالبا ما يتم إعادة التحليل لفقرات الاختبار باستخدام نموذج آخر وخاصة إذا كان النموذج المستخدم (ثلاثي المعلمات) .

لنعد إلى الحالة المتضمنة في الشكل (٣ – ٢) فقد بلغت قيمة (كا²) الملاحظة (٢٨ ٨٨) وأما القيمة المحسوبة فهي (١٩ ٤٥) وحيث أن القيمة المحسوبة أكبر من القيمة الملاحظة أو الجدولية فان منحنى خصائص الفقرة التي تتمتع بمعاملات (ص = ٠.٣٩ و ت= ١.٢٧) يتطابق مع النموذج ثلاثي المعلمات، ولسوء الحظ فان

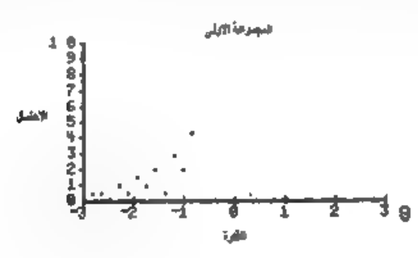
التحليلات التي تتم من خلال الحاسب الآلي لا تزودنا بتطابق جميع الفقرات، ولمزيد من التفصيل حول هذا الموضوع يمكن الرجوع إلى (Wright and Sten,1979)

إن عملية التدوير عملية معقدة من الناحية الرياضية وتتطلب جهدا كبيرا، ولذلك فان انتشار وشيوع النظرية الحديثة في القياس لم يبدأ بشكل واضح إلا بعد اختراع الحاسوب واستخداماته، ولذلك لا نهتم بالتفاصيل الدقيقة بسبب إن الحسابات تتم بشكل كمبيوتري، لكن الأهم من ذلك هو الاستدلال على كيفية التطابق وتحديد معلمات مختلف الفقرات

المجموعة (الثابتة) المشتركة لمعلمات الفقرة :

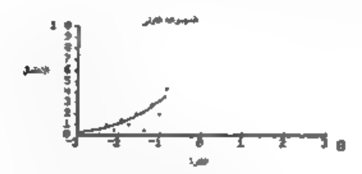
من الميزات المهمة للنظرية الحديثة في القياس أن معالم الفقرة لا تعتمد على مستوى القدرة للمفحوصين الذين يستجيبون لتلك الفقرة، ولذلك فان معالم الفقرة تعرف من خلال المجموعة الثابتة، وهذا يمكن توضيحه على النحو التالي:

المفحوصين حيث تمتلك المجموعة الأولى قدرة ضمن المدى (- ٣ ـــ -١) بمتوسط قدرة (-٢) والمجموعة الثانية لها قدرة ضمن المدى (١ ـــ ٣) بمتوسط قدره (٢)، أما الاحتمال الملاحظ للإجابة الصحيحة فيحسب من بيانات الاستجابة لكل مستويات القدرة في كلا المجموعتين وقد تم رسم احتمالات الإجابة الصحيحة للمجموعة الأولى شكل (٢١)



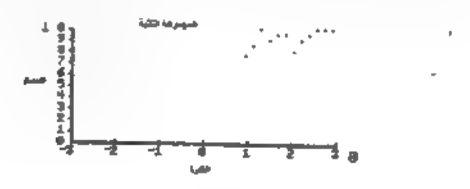
شكل(٢١) شكل احتمالات الإجابة الصحيحة لمجموعتين مختلفتين.

وتستخدم عادة الإجراءات الخاصة بأعلى نسبة ترجيح (maximum likelihood) لفحص فيما إذا كان منحنى خصائص الفقرة يطابق مع البيانات الناتجة من تقديرات معلمة الصعوبة (ص= -٠.٣٩ وت = ١.٢٧) بعد ذلك يتم تحديد منحنى خصائص الفقرة من خلال هذه التقديرات ويتم الرسم على مدى متصل القدرة وذلك للمجموعة الأولى كما في الشكل (٢٢)

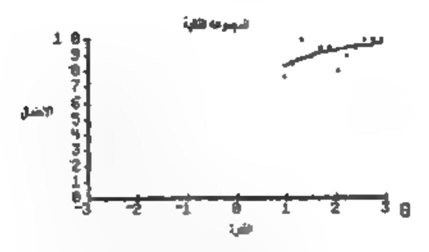


شكل (٢٢) منحنى خصائص الفقرة المطابق مع بيانات المجموعة الأولى .

لقد تم إعادة هذه العملية (مطابقة الفقرة) للمجموعة الثانية حيث تظهر احتمالات الإجابات الصحيحة على هذه الفقرة كما في شكل (٢٣) والذي يظهر مطابقة منحنى خصائص الفقرة مع النموذج حيث (ص= -٠.٣٩، ت = ١.٢٧) ويظهر ذلك في شكل (٢٤)



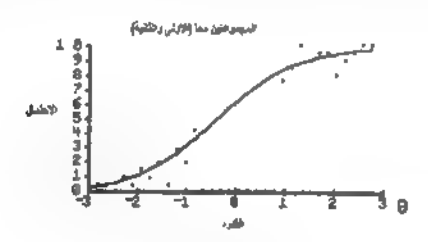
شكل (٢٣) الاحتمالات الملاحظة للإجابة الصحيحة للمجموعة الثانية.



شكل (٢٤) منحنى خصائص الفقرة المطابق مع بيانات المجموعة الثانية .

إن النتائج التي حصلنا عليها في ظل الخصائص السابقة للفقرة حيث (ص١ = ص٢) وكذلك (ث١ = ث٢) تبين أننا حصلنا على قيم متساوية لمعلمي الصعوبة والتمييز، وهذا يعني أن الفقرة كانت بمثابة مجموعة ثابتة، لكن هذه النتيجة لا تحدث دائما وان صدق النتائج يمكن تحقيقه بسهولة من خلال عمليات المطابقة لمنحنى خصائص الفقرة بقيم الاحتمالات الملاحظة للإجابة الصحيحة، وبما أن المجموعة الأولى لديها متوسط منخفض (-٢) فان مستويات القدرة في هذه المجموعة (الأولى) تشمل فقط جزء من المنحنى وهو في هذه الحالة يمثل الذيل الأيسر أي منطقة التناقص وهكذا فان الاحتمالات الملاحظة للإجابة الصحيحة سوف تمتد ضمن الحد الأدنى للمستوى (القيمة) المتوسطة، وعندما يتطابق المنحنى مع هذه البيانات فان الذيل الأيسر (المتناقص) هو الذي يحقق هذا التطابق .وعلى سبيل المثال في الشكل (٣ ـ ٥) وبما أن المجموعة الثانية لها متوسط قدرة(٢) فان الاحتمالات الملاحظة للإجابة الصحيحة سوف تمتد من القدرة المتوسطة حتى تقترب من الواحد الصحيح (١) وإذا ما حاولنا مطابقة منحنى خصائص الفقرة مع هذه البيانات فان القسم العلوي من المنحنى هو الذي سيحقق هذا التطابق كما هو واضح في الشكل (٢٤) وبما أن نفس الفقرة طبقت على كلا المجموعتين وتمت مطابقة خصائص الفقرة لكل منهما فان كلا من المنحنيين السابقين (الذيل الأيسر والذيل الأيمن) سيقعا تحت نفس منحنى

خصائص الفقرة، وهذا يعني أننا ستحصل على نفس قيم معلمات الفقرة في كلا الحالتين والشكل (٢٥) يجمع بين المجموعة الثانية لمعلمات الفقرة وهذه ميزة قوية للنظرية الحديثة، حيث تؤكد على أن قيم معلمات الفقرة تمثل خصائص الفقرة وليس خصائص المجموعة التي ستستجيب عليها، على عكس ما هو في النظرية الكلاسيكية، حيث أن صعوبة الفقرة هي نسبة من أجابوا على الفقرة إجابة صحيحة فإذا كان معامل الصعوبة لفقرة ما هو (صفر) فان عددا قليل جدا من المفحوصين سيجيب على الفقرة إجابة صحيحة من المفحوصين من المستوى المتدني، وان معامل الصعوبة سيكون عاليا بالنسبة للمجموعة ذاتها (المتدنية)، وإذا ما تم تطبيق هذه الفقرة على مجموعة من مستوى عالي القدرة فان معظم المفحوصين سيجيبوا عليها إجابة صحيحة وليكن معامل الصعوبة (ص= ٨ ٠)



شكل (٢٥) منحنى خصائص الفقرة المتطابق مع بيانات المجموعتين معا .

ومن الشكل نلاحظ أن الفقرة أبرزت مستويين مختلفين من الصعوبة وذلك باختلاف مستوى قدرة المفحوصين، بينما تبقى صعوبة الفقرة ثابتة مهما اختلف مستوى قدرة المفحوصين .لكن الذي يختلف هو احتمال الإجابة الصحيحة عليها، لكن من المهم ملاحظة أن الحالة التي نتحدث عنها (المجموعة الثانية) لا تحدث لجميع الفقرات، إنما يكون حدوثها قليل ومحكوم بعدة شروط من أهمها . أن المجموعتين يجب

أن يمثلا مستويين متناقضين من القدرة، إضافة إلى أن الفقرة يجب أن تقيس نفس السمة كما يعتمد ذلك على حجم العينة، وهذا ما يؤكد أن معالم الفقرة تعتمد على توزيع قدرة المفحوصين على متصل القدرة

خلاصة

- في النموذج ثلاثي المعلمات فان منحنى خصائص الفقرة يعتمد على تقدير معلمات الفقرة، حيث تستخدم في الكشف عن مطابقة الاحتمالات الملاحظة للاستجابة الصحيحة على الفقرة مع النموذج، وفي هذا النموذج فان بعض الفقرات عادة ما تبدي تطابقا بين تقدير المعلمات للفقرات ومنحنى خصائص الفقرة، وفي الحالات التي يكون التطابق فيها قليلا أو غير تام لبعض الفقرات فيكفي مراجعة الفقرات أو إعادة صياغتها بحيث تبدو أكثر دقة.
- إذا ستجابت مجموعتين لنفس الفقرة فيمكن إن تتطابق منحنيات خصائص الفقرة بغض النظر عن مدى القدرة الذي تقعا فيه أو تمتلكاه.
- في ظاهرة المجموعة الثابتة ليس المهم توزيع المفحوصين على متصل القدرة المهم هو تحديد مستويات القدرة، كما أن عدد المفحوصين في المجموعة لا يؤثر في هذه الظاهرة
- إذا توزعت مجموعتين من المفحوصين على طول متصل القدرة وكانت الفقرة مرجبة التمييز، فان المجموعة ذات مستوى القدرة المتدني ستتمثل (تظهر) في الجزء (الذيل) الأيسر السفلي لمنحنى الخصائص والمجموعة ذات مستوى القدرة العالي ستتمثل(تظهر) في الجزء (الذيل) العلوي الأيمن من المنحنى، فقد تحدث ظاهرة المجموعة الثابتة قد تحدث سواء كانت مستويات القدرة متداخلة أو غير متداخلة حيث أن التداخل ليس له اعتبار .
- إن التمييز بين المجموعتين (١، ٢) على اعتبار أن احدهما عالية القدرة والأخرى متدنية القدرة أي ان أحدهما تقع ضمن الحد الأعلى والأخرى ضمن الحد الأدنى للقدرة فان تصنيفهما المسبق لن يؤثر على موقعهما لان ذلك سيتم معرفته من خلال استجابة المفحوصين في كلا المجموعتين على الفقرات.

- غالبا ما تحدث ظاهرة المجموعة الثابتة إذا تم استخدام النموذج ثلاثي المعلمات.
- من المهم إن ندرك أنه عند استخدام بيانات الاستجابات على الفقرات فان تقديرات المعلمات يجب أن تعكس ثابتات العينة، حيث أن تطبيق الاختبار على مجموعات مختلفة لا يعطي نفس العينة من الفقرات في كل الأوقات لاختلاف قدراتهم وبالتالي لا بد من اختلاف الفقرات لتتناسب مع هذه القدرات، وهذا لا يعني أن ظاهرة المجموعة الثابتة غير صادقة، لكن ذلك يعني إن هذه الظاهر من الصعب ملاحظتها في البيانات الواقعية .

الفصل الرابع

منحنى خصائص الاختبار

Test Characteristic Curve

تعتمد النظرية الحديثة في القياس على الفقرات مستقلة عن بعضها البعض، لذلك جاء الحديث في الفصول الثلاثة السابقة من هذا الكتاب عن الفقرات على أساس أنها مستقلة، أما في هذا الفصل فسيتم الحديث عن مرحلة سيتم التعامل فيها مع الفقرات مجتمعة كونها تشكل بمجموعها الاختبار كوحدة واحدة

من المعروف أنه وبعد الانتهاء من تطبيق الاختبار تتم عملية التصحيح بحيث تأخذ الاستجابة على كل فقرة إحدى قيمتين الأولى العلامة (١) في حال كانت الإجابة على الفقرة صحيحة والعلامة (صفر) في حال كانت الإجابة خاطئة، وينطبق هذا الحديث على اختبار الاختيار من متعدد، حيث يتم الحصول على العلامة الكلية لأي مفحوص بجمع علامات كل فقرة أجاب عليها إجابة صحيحة بحيث تكون العلامة الكلية رقما صحيحا كما تقع هذه العلامة بين القيمة صفر وعدد الفقرات الاختبار(صفر – ن) حيث ن عدد الفقرات التي تشكل الاختبار، وإذا ما خضع المفحوص للاختبار عدة مرات وعلى فرض أنه لن يتذكر إجاباته في مرات التطبيق، فانه سيحصل على علامات مختلفة رأنه ومهما كان هذا الاختلاف فان العلامة ستدور حول قيمة معينة ستكون ممثلة للعلامات في مرات التطبيق التي خضع لها المفحوص وتسمى هذه العلامة بمتوسط العلامات، وفي نظرية السمات الكامنة فان هذه العلامة (المتوسط) تسمى العلامة الحقيقية (True Score) وتعرّف رياضيا حسب المعادلة (٥٢) والتي تفيد بأن العلامة الحقيقية عبارة عن مجموع احتمالات الإجابة الصحيحة على فقرات الاختبار عند مستوى قدرة معين.

دار جرير للنشر والتوزيع

ع = مجموع ح ق ....(٥٢)

حيث:

ع: العلامة الحقيقية عند مستوى معين من القدرة.

ح ق: احتمال الإجابة الصحيحة لفقرة ما.

ف: فقرة ما وعددها من (١ عدد فقرات الاختبار)، (أ- ن)

ن : عدد فقرات الاختبار

إن المهمة هنا هي حساب العلامة الحقيقية للمفحوصين عند مستوى القدرة، ولتوضيح ذلك سنقوم بحل المثال التالي

مثال:

اختبار يتكون من (٤) فقرات أحسب علامة المفحوصين من ذوي مستوى القدرة (ق = ١) وذلك باستخدام معادلة النموذج ثنائي المعلمات، إذا كانت معلمات فقرات الاختبار كما يلي ؟

الرقم	معامل الصعوبة	معامل التمييز
١	-١	٠.٥
٢	٠.٧٥	١.٢
٣	صفر	٠.٨
٤	٠.٥	١

الحل:

أولا : نحسب احتمال الإجابة الصحيحة على الفقرات .

احتمال الإجابة الصحيحة حسب النموذج ثنائي المعلمات يتم حسابه وفقا للمعادلة

ح (ق) = ١/ ١+ ث ^-حاق-ص)^ ٢ـ

بالنسبة للفقرة الأولى (ص = (-١، ت = ٠.٥).

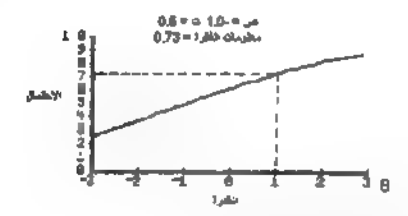
ح (ق) = ١/ ١+ ث ^-حاق-ص)^

ح (ق) = ١/ ١+ ث ^(١--١-)٠.٥-^

ح (ق) = ١/ ١+ ث ^-١^ ومنها ح (ق) = ١/ ١+ ٢.٧١٨ ^-١^

ح (ق) = ١/ ١+ ٠.٣٦ ومنها ح (ق) = ١/ ٢.٣٦ ومنها ح (ق) = ٠.٧٣

وإذا تم رسم منحنى خصائص الفقرة فسيكونكما هو مبين في الشكل رقم (٢٦)



شكل (٢٦) منحنى خصائص الفقرة الأولى.

بالنسبة للفقرة الثانية (ص = (٠.٧٥ ، ت = ١.٢)

ح (ق) = ١/ ١+ ٢.٧١٨ ^-(٠.٧٥-١)١.٢^

ح (ق) = ١/ ١+ ث ^-(٠.٢٥)١.٢^

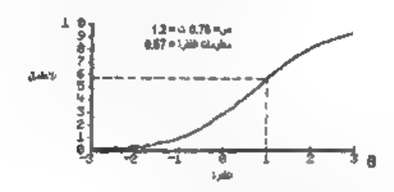
ح (ق) = ١/ ١+ ث ^-(٠.٢٥)١.٢^

ح (هـ) = ١/ ١+ ث $^{-٠.٣}$

ح (هـ) = ١/ ١+ ٢.٧١٨ $^{-٠.٣}$

ح (هـ) = ١/ ١+ ٠.٧٤٠ ومنها ح (هـ) = ١/ ١.٧٤٠ ومنها ح (هـ) = ٠.٥٧

وإذا تم رسم منحنى خصائص الفقرة فسيكونكما هو مبين في الشكل رقم (٢٧)



شكل (٢٧) منحنى خصائص الفقرة الثانية.

بالنسبة للفقرة الثالثة (ص = (صفر، ت = ٠.٨)

ح (هـ) = ١/ ١+ ث $^{-٠.٨(١-صفر)}$

ح (هـ) = ١/ ١+ ث $^{-٠.٨(١ ٠)}$

ح (هـ) = ١/ ١+ ث $^{-٠.٨}$

ح (هـ) = ١/ ١+ ٢.٧١٨ $^{-٠.٨}$

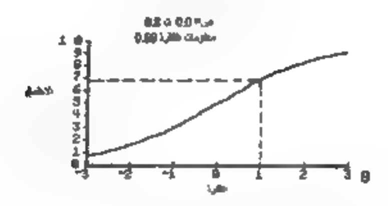
ح (هـ) = ١/ ١+ ٠.٤٤٩

ح (هـ) = ١/ ١.٤٤٩

ح (هـ) = ٠.٦٩

وإذا تم رسم منحنى خصائص الفقرة فسيكونكما هو مبين في الشكل (٢٨)





شكل (٢٨) منحنى خصائص الفقرة الثالثة.

بالنسبة للفقرة الرابعة (ص = (٥ ٠، ت = ١)

ح (ق) = ١/ ١+ ث – ١(١– ٠.٥)

ح (ق) = ١/ ١+ ث – ١(٠.٥)

ح (ق) = ١/ ١+ ث (– ٠.٥)

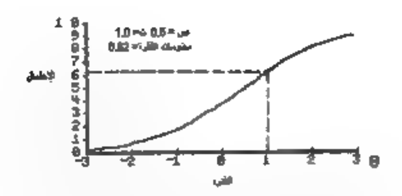
ح (ق) = ١/ ١+ ٢.٧١٨ – ٠.٥

ح (ق) = ١/ ١+ ٠.٦٠٦

ح (ق) = ١/ ١+ ١.٦٠٦

∴ ح (ق) = ٠.٦٢

وإذا تم رسم منحنى خصائص الفقرة فسيكونكما هو مبين في الشكل (٢٩)



شكل (٢٩) منحنى خصائص الفقرة الرابعة.

والآن يمكن إيجاد العلامة الحقيقية لمفحوص من ذوي مستوى القدرة (ق = ١) وذلك بجمع احتمالات الإجابة الصحيحة على الفقرات الأربع حسب المعادلة (٤ ـ١) وذلك على النحو التالي

العلامة الحقيقية = مجموع احتمالات الإجابة الصحيحة على فقرات الاختبار

ع ح = ح(١) + ح(٢) + ح(٣) + ح(٤)

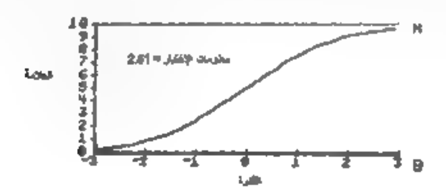
= ٠.٧٣ + ٠.٥٧ + ٠.٦٩ + ٠.٦٢

= ٢.٦١

وحسب ذلك فان المفحوصين من ذوي مستوى القدرة (ق = ١) سيحصلون على العلامة (٢.٦١) من (٤) وهي العلامة الكلية للاختبار كونه يتكون من (٤) فقرات، وهذه العلامة منطقية ومن البديهي أن يحصل عليها المفحوص عند هذا المستوى حيث تتقاطع منحنيات خصائص الفقرات الأربع مع الخط العامودي من متصل القدرة عند احتمال الإجابة أكثر من (٥ ٠)، ولذلك فان مجموع احتمالات الإجابة الصحيحة على الفقرات الأربع عالي، لكن من الناحية العملية لن يحصل أي من المفحوصين على العلامة (٢.٦١)، والسبب في ذلك هو أن هذه العلامة هي متوسط نظري لجميع العلامات التي يمكن أن يحصل عليها المفحوصين عند هذا

المستوى من القدرة (ق = ١) إذا خضعوا لهذا الاختبار عدة مرات، أي انه متوسط نظري. وانه مهما كان الاختبار طويلا فان العلامة تحسب بهذه الطريقة .

ولا يفوتنا أن نتذكر أن الإجراءات الحسابية التي تمت سابقا عند مستوى واحد من مستويات القدرة الممتدة ضمن المدى النظري (- ∞ ـــ ∞) وعند تطبيق معادلة العلامة الحقيقية لجميع مستويات القدرة وتمثيلها على محور السينات الذي يمثل مستويات القدرة ومحور الصادات الذي يمثل العلامة الحقيقية فسينتج لدينا منحنى تراكمي للفقرات ويسمى بمنحنى خصائص الاختبار (Test Characteristic Curve) كالذي يظهر في الشكل (٣٠) والذي يمثل منحنى خصائص لاختبار يتكون من (١٠) فقرات



شكل (٣٠) منحنى الخصائص لاختبار يتكون من ١٠ فقرات.

ومن الضروري أن يدرك القارئ أن منحنى خصائص الاختبار هو علاقة بين العلامة الحقيقية ومتصل القدرة، ومن خلال هذا المنحنى فانه يمكن أن نجد العلامة الحقيقية عند أي مستوى من مستويات القدرة، فعلى سبيل المثال فانه يمكن معرفة العلامة الحقيقية لمستوى القدرة (ق =١) من خلال الشكل (٤ ـــ ٥) برسم خط عامودي صاعد من النقطة التي تمثل مستوى القدرة (ق = ١) على محور السينات حتى يتقاطع مع المنحنى وكذلك رسم خط أفقي من نقطة التقاطع مع المنحنى باتجاه محور الصادات، حيث ستكون العلامة التي تمثل نقطة التقاطع مع محور الصادات هي

العلامة الحقيقية التي تقابل مستوى القدرة (ق = ١) وهي العلامة (٧.٨) في الشكل أعلاه، وهذا يعني أن كل مفحوص يتمتع بمستوى القدرة (ق = ١) سيحصل على العلامة (٧.٨) من العلامة الكلية والتي تبلغ في هذا المثال (١٠) لأنها تمثل مجموع احتمالات إجابات الطلبة على فقرات هذا الاختبار.كما وأنه من المهم أن ندرك انه إذا استخدم النموذج أحادي وثنائي المعلمات لعدد (ن) من الفقرات فان الذيل الأيسر من منحنى خصائص الاختبار سيمثل الحدود الدنيا للعلامات عند مستويات القدرة التي تقترب من الصفر وأما الذيل الأيمن من المنحنى فسيمثل الحدود العليا للعلامات عند مستويات القدرة التي تقترب من (ق = ١)

أما في حال استخدم النموذج ثلاثي المعلمات لاختبار يتكون من عدد (ن) من الفقرات فان الذيل الأيسر من منحنى خصائص الاختبار سيمثل مجموع معلمات التخمين (خ) التي تزيد عن الصفر وهذا يعني الادعاء الذي يتعلق بالنموذج ثلاثي المعلمات والذي مفاده أن المفحوصين ذوي المستويات المتدنية من القدرة يمكنهم وبسهولة أن يحصلوا على علامة عالية من خلال التخمين، أما الذيل الأيمن من منحنى خصائص الاختبار فانه سيبدأ بالتصاعد حسب عدد الفقرات التي يتكون منها الاختبار، لذلك فان العلامة الحقيقية لمجموعة (ن) من الاستجابات ستبلغ أعلى ما يمكن مع زيادة مستوى القدرة.

إن الدور الأساسي لمنحنى خصائص الاختبار حسب النظرية الحديثة هو تقديم معاني منطقية لتحويل علامات القدرة إلى علامات حقيقية، ويتبين ذلك من خلال التركيز على الجوانب العملية التي قد لا يستطيع مستخدم الاختبار من تفسيرها كالقدرة أو العلامة المعبرة عنها، حيث أن مستخدم الاختبار وعندما يقوم بتحويل العلامات التي تعبر عن القدرة إلى علامات حقيقية، فانه يقوم بإعطاء رقم يتعلق بعدد الفقرات التي يتكون منها الاختبار بحيث يكون هذا الرقم إطارا مرجعيا يمكن لمستخدم الاختبار تفسيره، ومهما كان هذا الرقم مألوفا فانه لا يمكن تفسيره بطريقة مباشرة.حيث يلعب منحنى خصائص الاختبار دورا مهما في إجراءات معايرة (Calibration) الاختبارات .

إن تفحص منحنى خصائص الاختبار يلاحظ انه عبارة عن يعبر عن معادلة متزايدة وفي بعض الحالات يكون أكثر انبساطا من شكل الحرف (S) كما في منحنى خصائص الفقرة، ويكون في حالات أخرى منبسط ومتزايد اي منبسط ويميل إلى الثبات قبل أن يرتفع مرة أخرى، وفي كل الأحوال فان المنحنى يبلغ أعلى قيمة له – وتساوي عدد فقرات الاختبار – وذلك عند ذيله العلوي ومن الجدير بالذكر فان منحنى خصائص الاختبار يعتمد على عدة عوامل بما في ذلك عدد الفقرات التي يتكون منها ؟، إلى جانب النموذج اللوغريتمي المستخدم في رسم شكل منحنى خصائص الفقرة، وكذلك قيم معلمات الفقرة، ومن هنا فانه لا يوجد معادلة واحدة ومحددة لمنحنى خصائص الاختبار حيث أن هناك صيغا أخرى غير تلك التي استخدمت في هذا الفصل وهي معادلة (٤ ــ ١) ولذلك فان أفضل طريقة لصياغة أو رسم منحنى خصائص الاختبار هو تحديد احتمالات الإجابة الصحيحة لفقرات الاختبار عند كل مستوى من مستويات القدرة اعتمادا على النموذج اللوغريتمي المستخدم.وإذا تم الحصول عليها فإننا سنحصل على منحنى خصائص الاختبار بمجرد جمع هذه الاحتمالات، ومن الأمور المهمة في هذا المجال ضرورة الانتباه الى أن نمط منحنى الاختبار لا يعتمد على توزيع التكرار للمفحوصين بالنسبة لعلامات القدرة على متصل القدرة حيث سيتشابه كل من منحنى خصائص الاختبار والفقرة على حد سواء لان كلا منهما سيمثل العلاقة بين مقياسين ولا يعتمد على توزيع العلامات على كل منهما

يمكن تفسير منحنى خصائص الاختبار بنفس الطريقة التي يفسر بها منحنى خصائص الفقرة إلى حد كبير حيث أن مستوى القدرة متطابق عند العلامة الحقيقية التي تمثل الوسط وكأن الفقرات مقسمة إلى قسمين أي (ن/ ٢) وان الميل العام لمنحنى خصائص الاختبار له علاقة بقيمة العلامة الحقيقية والتي تعتمد على مستوى القدرة، كما يقترب منحنى خصائص الاختبار من الخط المستقيم عند أعلى قيمة على متصل القدرة، وفي معظم الاختبارات فان منحنى خصائص الاختبار لا يكون خطيا ويكون الميل معبرا عن انخفاض مستويات القدرة، ومن هنا فلا يوجد معادلة واحدة لمنحنى

خصائص الاختبار لأنه كونه لا يوجد له معلمات، وتعتبر العلامة الحقيقية عند نقطة منتصف العلامات تعبيرا عن الصعوبة للاختبار أو الميل لمنحنى خصائص الاختبار حيث يعبر المنحنى عن نفسه، بمعنى أن شكل المنحنى في منطقة الوسط وميل المنحنى من موصفات ذلك المنحنى.

خلاصة

- ان توفر مستوى القدرة والعلامة الحقيقية التي تقابله والتي يمكن الحصول عليها توفر احتمالات الإجابة الصحيحة على الفقرات وهذا يمكننا من إيجاد شكل منحنى خصائص الاختبار.
- تعتبر كلا من مستوى القدرة والعلامة الحقيقية متغيرات متصلة
- عندما يتكون الاختبار من فقرة واحدة (ن =١) فان مدى العلامة الحقيقية يكون بين (صفر ــ ١) كما يكون شكل منحنى خصائص الاختبار مطابقا لمنحنى خصائص الفقرة
- لا يكون منحنى خصائص الفقرة مطابقا لمنحنى خصائص الاختبار إذا كان عدد الفقرات أكثر من(١) فقد يكون منبسطا أو مفلطحا حيث تعكس المنحنيات قيم معلمات الفقرات
- إن مستويات القدرة عند المتصف (ن/ ٢) تعتمد على متوسط معلمات الصعوبة.
- إذا انحصرت قيم معاملات الصعوبة للفقرة في مدى ضيق على متصل القدرة فان انخفاض منحنى خصائص الاختبار يعتمد في الأساس على معدل معلمات التميز ت إما إذا توزعت معاملات الصعوبة على مدى واسع على متصل القدرة فان انخفاض منحنى خصائص الاختبار سيقل أي سيكون أكثر انبساطا من الوضع في الحالة السابقة، ويحدث هذا طالما كانت معاملات التميز متقاربة (ثابتة نسبيا)
- عند استخدام النموذج ثلاثي المعلمات فان الحد الأدنى للعلامات الحقيقية (أقل علامة) سيكون مجموع قيم معلمات التخمين لكل الفقرات المكونة للاختبار.
- إن شكل منحنى خصائص الاختبار يعتمد على عدد الفقرات المكونة للاختبار ونموذج منحنى خصائص الفقرة المستخدم وكذلك قيم معلمات فقرات الاختبار.

- من الممكن بناء أو الحصول على منحنى خصائص الاختبار، بحيث يكون متناقصا كلما زادت القدرة، وهنا لا بد من أن تكون الفقرات ذات تمييز سالب وهذا يعني أن الاختبار غير عملي او غير منطقي ولا يحقق الأهداف المرغوبة، لأن ذلك يعني أن المفحوصين من ذوي القدرة العالية سيحصلون على علامات أقل من علامات المفحوصين من ذوي مستويات القدرة المتوسطة أو المتدنية . أي أنه يميز ولكن التمييز ليس بالتمييز المرغوب، وقد يكون ذلك بسبب التخمين أو غموض الفقرات أو خطأ في مفتاح التصحيح .

الفصل الخامس

تقدير قدرة المفحوصين

Estimating an Examines Ability

يتمثل الغرض الأساسي للاختبار في ظل نظرية السمات الكامنة (IRT) هو تحديد قدرة المفحوص على متصل القدرة، وإذا تم ذلك فانه يمكن تحقيق هدفين أساسيين . يتعلق **الأول** بتحديد القدرة التي يمتلكها المفحوص ويتعلق **الثاني** بمقارنة المفحوصين ببعضهم البعض لأغراض رصد الدرجات والتصاديق والمنح المدرسية، ومن هنا فان التركيز في هذا الفصل سوف يكون على إجراءات تقدير قدرة المفحوصين .

من المعروف أن الاختبار يستخدم لقياس سمة كامنة من خلال مجموعة أو عينة (ن) من الفقرات التي تقيس نفس السمة أو المؤشرات الدالة عليها، وحيث كان التركيز سابقا على تقدير معلمات الفقرة تم افتراض أن قدرة المفحوصين معروفة، الأمر الذي يمكننا من تقدير القدرة الحقيقية غير المعروفة، وهنا تبرز أهمية هذا الافتراض وذلك لان وحدة القياس للقدرة هي نفس وحدة القياس لمعلمة صعوبة الفقرة، فعندما يتم تطبيق الاختبار فان كل مفحوص يستجيب لعدد (ن) من الفقرات حيث تكون العلامة التي سيحصل عليها الطالب على الفقرة الواحدة ثنائية إما العلامة (١) في حالة الإجابة الصحيحة أو العلامة (صفر) في حالة الإجابة الخاطئة، وتبعا لذلك فان علامة المفحوص على الفقرة الواحدة ستنحصر بين المدى (صفر ــــ ١) أي أن استجابة المفحوص ستأخذ نمطين إما (١) وإما (صفر) ومن هنا سيتم استخدام كل من نمط الاستجابة ومعلمات الفقرة في تقدير قدرة المفحوصين والتي لا تكون عادة معروفة.

إجراءات تقدير القدرة.

لتقدير قدرة المفحوص في ظل النظرية الحديثة للقياس يتم استخدام إجراءات الحد الأعلى لنسبة الترجيح (maximum likelihood) كما هو بالنسبة لتقدير معلمات الفقرة، حيث تتفاعل وتتداخل هذه الإجراءات فيما بينها وتبدأ بالقيم المسبقة (المحسوبة سابقا) مثل الصعوبة والتمييز باعتبارها معلمات للفقرة، حيث تستخدم لحساب احتمال إجابة المفحوص على كل فقرة إجابة صحيحة، بعد ذلك تجري عملية تعديل (Adjustment)على تقدير القدرة بحيث يتطابق نمط الاستجابة للفقرة مع احتمالات الإجابة الصحيحة والتي تكون قد تم حسابها، وتستمر هذه التعديلات حتى نحصل على أقل قيمة مضافة بحيث تكون القيمة المضافة قيما صغيرة حيث تكون النتيجة التي نحصل عليها بمثابة تقدير لمعلمة القدرة، ومن المهم إن يدرك القارئ إن هذه الإجراءات (التعديل) تتم لكل مفحوص خضع للاختبار، وهذا ما سيتضح في الفصول اللاحقة، وعلى أية حال فان هذه الإجراءات تعتمد على مجموعة من المعالجات، ولكل مفحوص، ولذلك فان المهم هنا هو كيفية تحديد أو تقدير قدرة المفحوص الواحد وتتم عملية تقدير القدرة وفقا للمعادلة (٥٢) ولتوضيح إجراءات التقدير لقدرة المفحوصين سيتم حل المثال أدناه

$$ق_{ر+1} = ق_{ر} + \frac{\text{مجموع} - ت_{س}\left(ن_{م} - ح_{س}(ق_{ر})\right)}{\text{مجموع } ت^{2}_{ر}\left(ح_{س}(ق_{ر})\right)\left(ح_{ع}(ق_{ر})\right)} \quad \ldots\ldots\ldots(٥٢)$$

حيث .

$ق_{ر+1}$: قدرة المفحوص ر

$ت_{س}$: معامل تمييز الفقرة .

$ن_{م}$: نمط استجابة المفحوص على الفقرة ويكون إما (١) وإما (صفر).



ح س(ق ن) : احتمال الإجابة الصحيحة على الفقرة تحت نموذج لوغريتمي ما وعند مستوى قدرة محدد.

ح ع(ق ن) : احتمال الإجابة الخاطئة على الفقرة تحت نموذج لوغريتمي ما و مستوى قدرة محدد ويساوي (1 – ح ع).

مثال (١) :

كانت خصائص (معلمات) الفقرات في اختبار مكون من ثلاث فقرات كما كانت استجابات المفحوصين على هذه الفقرات كما هي مبينة أدناه، ما هي إجراءات تقدير لقدرة في ظل النموذج ثلاثي المعلمات؟

الفقرة	الصعوبة	التمييز	نمط الاستجابة
١	ص ١ = -١	ت ١ = ١	ن م ١ = ١
٢	ص ٢ = صفر	ت ٢ = ١.٢	ن م ٢ = صفر
٣	ص ٣ = ١	ت ٣ = ٠.٨	ن م ٣ = ٣

الحل:

للحل سيتم تحديد القيم المطلوبة في معادلة القدرة لكل فقرة من الفقرات الثلاث، وهنا سنلجأ إلى الصيغة العامة للنموذج ثنائي المعلمات لإيجاد قيمة الاحتمال الإجابة الصحيحة على الفقرات ح س(ق) وفي ظل المعطيات الخاصة بكل منها بالتفصيل·ثم سنورد القيم بالنسبة للفقرات الأخرى كما في الخطوات التالية.

الفلترة الأولى First Iteration

أولا : بالنسبة للفقرة الأولى وعند مستوى القدرة ق =١ ، ص= -١، ت = ١، ن م = ١ فان إجراءات التدوير الأول تتم على النحو التالي:

$$ح_{ص1} = \frac{1}{1 + ت^{-ت(ق-ص)}} =$$

$$ح_{ص1} = \frac{1}{1 + 2.718^{-1(1-0)}}$$

$$ح_{ص1} = \frac{1}{1 + 2.718^{-2}}$$

$$ح_{ص1} = \frac{1}{1 + 0.13} = 0.88$$

وهنا يمكن أن نجد القيم المطلوبة في الصيغة العامة للمعادلة الخاصة بتقدير القدرة ومنها احتمال الإجابة الخاطئة ح (ع) وهي متممة احتمال الإجابة الصحيحة وكذلك القيمة ت س (ل م – ح ص(ق)) والقيمة $ت^2$ ر ح (ص) ح ع (ع) باعتبارها أطراف المعادلة (5-1) وذلك على النحو التالي

$$ح_ع = 1 - ح_ص = 1 - 0.88 = 0.12$$

$$ت_س (ل_م - ح_{ص(ق)}) = 1(1 - 0.88)\ 1\ (0.12) = 0.12$$

$$ت^2_ر\ ح(ص)\ ح_ع(ع) = 1^2\ (0.88)(0.12) = 1(0.105) = 0.105$$

ثانيا : بالنسبة للفقرة الثانية وعند مستوى القدرة ق =1، ص= 0 ، ت = 1.2 ل م = 0 فإن إجراءات التدوير الأول على النحو التالي.

$$ح_ص = \frac{1}{1 + ت - ت(ق - ص)}$$

$$ح_{ص2} = \frac{1}{1 + 2.718 - 1.2(1 - صفر)}$$

ح ص٢ = ١ / (١ + ٢.٧١٨ − ١.٢)

ح ص٢ = ١ / (١ + ٠.٣٠١) = ٠.٧٧

وكما تم في الخطوة السابقة ستجد القيم ح (ع) وكذلك القيمة ت س (ن م − ح ص(ق ر)) والقيمة ت٢ ر ح (س) ح ع (ع) وذلك على النحو التالي

ح ع = ١ − ح ص = ١ − ٠.٧٧ = ٠.٢٣

ت س (ن م − ح ص(ق ر)) = ١.٢ (صفر − ٠.٧٧) = ١.٢ (٠.٢٣) = ٠.٩٢٤

ت٢ ر ح (س) ح ع (ع) = ١.٢ ٢ (٠.٧٧)(٠.٢٣) = ١.٤٤ (٠.١٧٧) = ٠.٢٥٥

ثالثا : بالنسبة للفقرة الثالثة وعند مستوى القدرة ق =١، ص= ١، ت = ٨ ٠ ن م = ١ فان إجراءات التدوير الأول تتم على النحو التالي :

ح ص١ = ١ / (١ + ت − ت(ق − ص)) =

ح ص٣ = ١ / (١ + ٢.٧١٨ − ٠.٨(١ − ١))

ح ص٣ = ١ / (١ + ٢.٧١٨ (صفر))

ح ص٣ = ١ / (١ + ١) = ٠.٥٠

وكما تم في الخطوة السابقة أيضا سنجد القيم ذات العلاقة وهي ح ع والقيمة ث س (ن م − ح س(ق ن)) والقيمة ث²ر ح (س) ح ع (ع) وذلك على النحو التالي

ح ع = ١ − ح س = ١ − ٠,٥٠ = ٠,٥٠

ث س (ن م − ح س(ق ن)) = ٠,٨(١ − ٠,٥٠) ٠,٨ (٠,٥٠) = ٠,٤٠

ث²ر ح (س) ح ع (ع) = ٠,٨ ²(٠,٥٠)(٠,٥٠) = ٠,٦٤(٠,٢٥) = ٠,١٦

والآن يمكن تلخيص البيانات التي حصلنا عليها كما في الجدول رقم (٩) وذلك من أجل التسهيل على القارئ استخدام هذه البيانات ولتقدير قدرة المفحوصين عند مستوى القدرة (١) نطبق الصيغة العامة لتقدير معلمة القدرة

جدول (٩) : القيم الخاصة بتقدير قدرة المفحوصين في الفقرة الأولى للمثال (١)

الفقرة	ن م	ح س	ح ع	ث(ن م − ح م)	ث²(ح س● ح ع)
١	١	٠,٨٨	٠,١٢	٠,١١٩	٠,١٠٥
٢	٠	٠,٧٧	٠,٢٣	٠,٩٢٢-	٠,٢٥٥
٣	١	٠,٥٠	٠,٥٠	٠,٤٠٠	٠,١٦٠
المجموع				٠,٤٠٣-	٠,٥٢٠

من الملاحظ من الصيغة العامة لمعادلة تقدير معلمة القدرة أنها تتمثل في مجموع القيم الواردة في الجدول أي مجموع العامود ث(ن م − ح ع) والذي يمثل بسط المعادلة والعامود ث²(ح س● ح ع والذي يمثل مقام المعادلة أي أن

$$\text{ق}_{\text{ر}+1} = \text{ق}_{\text{ر}} + \frac{\text{مجموع ث س}\left(\text{ن م} - \text{ح س(ق ن)}\right)}{\text{مجموع ث}^2_{\text{ر}}\left(\text{ح س (ق ن)}\right)\left(\text{ح ع (ق ن)}\right)}$$

$$\text{ق}_{\text{ر}+1} = ١ + \frac{(-٠,٤٠٣)}{(٠,٥٢٠)} = ١ + (-٠,٧٧٣) = ١ - ٠,٧٧٣ = ٠,٢٢٧$$

وهذا يمثل قدرة المفحوصين عند مستوى القدرة (ق = ١)، ومن الملاحظ ان قيمة معلمة القدرة للمفحوصين والتي بلغت (٢٢٧ ٠) بعيدة عن مستوى القدرة النظري أي (ق =١) لذلك فنحن بحاجة إلى عملية تدوير أو فلترة ثانية للحصول على أعلى قيمة مقدرة لمعلمة القدرة في هذا المستوى ، حيث ستحل القيمة المقدرة في مرة الفلترة الأولى مكان القيمة (ق = ١) وقد يتساءل القارئ عن عدد مرات الفلترة اللازمة لتقدير القدرة والإجابة تتمثل في انه لا يوجد عدد محدد من الفلترات والهدف من الفلترة هو الوصول إلى أدق قيمة معبرة عن قدرة المفحوصين، لذلك نتوقف عن الفلترة عندما نصل إلى ما يسمى بأقل قيمة مضافة، وهي الفرق بين أي قيمتي قدرة متتاليتين، بمعنى انه إذا كان الفرق بين قيمتي القدرة المحسوبتين من مرتي فلترة متتاليتين قيمة صغيرة جدا فيمكننا التوقف عن الفلترة، إضافة إلى انه يوجد محك للحكم على الحاجة إلى الفلترة وذلك من خلال الخطأ المعياري للتقدير (Standard Error of Estimate). والذي سنتحدث عنه فيما بعد، والآن سنتابع الفلترة باستخدام القيمة المقدرة لمعلمة القدرة في مرة الفلترة الأولى

الفلترة الثانية Second Iteration

أولا : بالنسبة للفقرة الأولى وعند مستوى القدرة ق =٠.٢٢٧ ، ص= -١، ت = ١ ن م = ١ فان إجراءات التدوير الثاني تتم على النحو التالي.

$$ح\,ص_1 = \frac{1}{1 + ت^{-م(ق-ص)}} =$$

$$ح\,ص_1 = \frac{1}{1 + ٢.٧١٨^{-1(٠.٢٢٧ - -١)}}$$

$$ح\,ص_1 = \frac{1}{1 + ٢.٧١٨^{-١.٢٢٧}}$$

$$ح\,ص_1 = \frac{1}{1 + ٠.٢٩٣} = ٠.٧٧٣$$

وهنا يمكن أن نجد القيم المطلوبة في الصيغة العامة للمعادلة الخاصة بتقدير القدرة ومنها احتمال الإجابة الخاطئة ح ع وهي متممة احتمال الإجابة الصحيحة وكذلك القيمة ث س (ن م − ح س(ق ر)) والقيمة ث٢ ر ح (س) ح ع (ع) وذلك على النحو التالي

ح ع = ١ − ح س = ١ − ٠.٧٧ = ٠.٢٣

ث س (ن م − ح س(ق ر)) = ١(١ − ٠.٧٧٣) ١ (٠.٢٢٧) = ٠.٢٢٧

ث٢ ر ح (س) ح ع (ع) = ١ ٢ (٠.٧٧٣)(٠.٢٢٧) = ١(٠.١٧٥) = ٠.١٧٥

ثانيا : بالنسبة للفقرة الثانية وعند مستوى القدرة ق =٠.٢٢٧ ، ص= صفر، ث = ١.٢

ن م = صفر فان إجراءات التدوير الثاني تتم على النحو التالي:

ح ص١ = ١ / (١ + ث − ث(ق − ص))

ح ص٢ = ١ / (١ + ٢.٧١٨ − ١.٢(٠.٢٢٧ − صفر))

ح ص٢ = ١ / (١ + ٢.٧١٨ − ٠.٢٧٢٤)

ح ص٢ = ١ / (١ + ٠.٧٦) = ٠.٥٧

وهنا يمكن أن نجد القيم المطلوبة في الصيغة العامة للمعادلة الخاصة بتقدير القدرة ومنها احتمال الإجابة الخاطئة ح (ع) وهي متممة احتمال الإجابة الصحيحة وكذلك القيمة ث س (ن م − ح س(ق ر)) والقيمة ث٢ ر ح (س) ح ع (ع) وذلك على النحو التالي

ح ع = ١ − ح س = ١ − ٠.٥٧ = ٠.٤٣

ث س (ن م − ح س(ق ر)) = ١.٢(صفر − ٠.٥٧) ٢ ١ (−٠.٥٧) = −٠.٦٨١

ث٢ ر ح (س) ح ع (ع) = ١.٢ ٢ (٠.٥٧)(٠.٤٣) = ١.٤٤ (٠.٢٤٥١) = ٠.٣٥٣

ثالثا : بالنسبة للفقرة الثالثة وعند مستوى القدرة ق =٠.٢٢٧، ص= ١، ت = ٠.٨ ن م = ١ فان إجراءات التدوير الثاني تتم على النحو التالي:

$$ح\,ص_1 = \frac{1}{1 + ت - ت(ق - ص)} =$$

$$ح\,ص_3 = \frac{1}{1 + 2.718 - 0.8(0.227 - 1)}$$

$$ح\,ص_3 = \frac{1}{1 + 2.718\,(0.6184)}$$

$$ح\,ص_3 = \frac{1}{1 + 1.85} = 0.35$$

وهنا يمكن أن نجد القيم المطلوبة في الصيغة العامة للمعادلة الخاصة بتقدير القدرة ومنها احتمال الإجابة الخاطئة ح ع وهي متممة احتمال الإجابة الصحيحة وكذلك القيمة $ت_س$ (ك م — ح ص(ك)) والقيمة $ت^2_ر$ ح (ص) ح ع (ع) وذلك على النحو التالي

$ح\,ع = 1 - ح\,ص = 1 - 0.35 = 0.65$

$ت_س$ (ك م — ح ص(ك)) = ٠.٨(١ — ٠.٣٥) ٠.٨ (٠.٦٥) = ٠.٥٢

$ت^2_ر$ ح (ص) ح ع (ع) = ٠.٨ ^{2}(٠.٣٥)(٠.٦٥) = ٠.٦٤ (٠.٢٢٧٥) = ٠.١٤٦

والآن يمكن تلخيص البيانات التي حصلنا عليها كما في الجدول رقم (١٠) أدناه لتقدير قدرة المفحوصين مع تذكر أن معلمة القدرة الذي سنستخدمها هي (ق =٠.٢٢٧) حيث سنطبق الصيغة العامة لتقدير معلمة القدرة .

دار جرير للنشر والتوزيع

جدول (١٠) : القيم الخاصة بتقدير قدرة المفحوصين في الفلترة الثانية للمثال (١)

الفقرة	ن م	ح ص	ح م	ت(ن م – ح م)	ت ٢(ح ص * ح م)
١	١	٠.٧٧	٠.٢٣	٠.٢٢٧	٠.١٧٥
٢	٠	٠.٥٧	٠.٤٣	٠.٦٨١-	٠.٣٥٣
٣	١	٠.٣٥	٠.٦٥	٠.٥٢	٠.١٤٦
المجموع				٠.٠٦٦	٠.٦٧٤

وهنا سنعيد ما تم في الفلترة الأولى أي سنقدر معلمة القدرة مع تذكر أن (ق = ٠.٢٢٧) وذلك على النحو التالي

ق ر+١ = ق ر + مجموع – ت س (ن م – ح ص(ق ر)) / مجموع ت٢ر (ح ص (ق ر)) (ح ح (ق ر))

ق ر+١ = ١ + (٠.٠٦٦) / (٠.٦٧٤) = ٠.٢٢٧ + (٠.٠٧٩) = ٠.٢٢٧ + ٠.٠٧٩ = ٠.٣٢٤

من الملاحظ ان قيمة معلمة القدرة المقدرة للمفحوصين هنا زادت عن القيمة في الفلترة الأولى وبفارق ملحوظ وطالما أننا حصلنا على قيمة مضافة كبيرة نوعا ما فالمتوقع أن نستمر في عملية الفلترة، مع الأخذ بعين الاعتبار أن قيمة القدرة المستخدمة في مرة الفلترة القادمة (الثالثة) ستكون قيمة القدرة المقدرة في المرة الثانية أي أن (ق = ٠.٣٢٤) وسنستخدم نفس الإجراءات.

الفلترة الثالثة : Third Iteration

أولا: بالنسبة للفقرة الأولى وعند مستوى القدرة ق =٠.٣٢٤ ، ص= -١، ت = ١ ن م = ١ فان إجراءات التدوير الثالث تتم على النحو التالي:

ح ص١ = ١ / (١ + ت – ت(ق – ص)) =



دار جرير للنشر والتوزيع

$$\text{ح ص١} = \frac{\text{١}}{\text{١ + ٢٫٧١٨ − ١(٠.٣٢٤ − − ١)}}$$

$$\text{ح ص١} = \frac{\text{١}}{\text{١ + ٢٫٧١٨ − ١٫٣٢٤}}$$

$$\text{ح ص١} = \frac{\text{١}}{\text{١ + ٠٫٢٦٦}} = \text{٠.٧٩}$$

وهنا يمكن أن نجد القيم المطلوبة في الصيغة العامة للمعادلة الخاصة بتقدير القدرة ومنها احتمال الإجابة الخاطئة (ح ع)حيث أنها متممة احتمال الإجابة الصحيحة وكذلك القيمة ت س (ن م − ح س(ق ن)) والقيمة (ت٢ د ح (س) ح ع (ع)) وذلك على النحو التالي

$$\text{ح ع = ١ − ح س = ١ − ٠.٧٩ = ٠.٢١}$$

$$\text{ت س (ن م − ح س(ق ن)) = ١(١ − ٠.٧٩) ١ (٠.٢١) = ٠٫٢١٠٢}$$

$$\text{ت}^{\text{٢}}\text{ د ح (س) ح ع (ع) = ١}^{\text{٢}}\text{ (٠٫٧٩)(٠.٢١) = ١(٠.١٦) = ٠.١٦٦}$$

ثانيا: بالنسبة للفقرة الثانية وعند مستوى القدرة ق =٠.٢٢٧، ص= صفر، ت = ١.٢ ن م = صفر فان إجراءات التدوير الثالث تتم على النحو التالي

$$\text{خ ص} = \frac{\text{١}}{\text{١ + ث − ت(ق − ص)}}$$

$$\text{ح ص٢} = \frac{\text{١}}{\text{١ + ٢٫٧١٨ − ٢ ١(٠٫٣٢٤ − صفر)}}$$

$$\text{ح ص٢} = \frac{\text{١}}{\text{١ + ٢.٧١٨ − ٠.٣٨٨}}$$

$$\text{ح ص٢} = \frac{\text{١}}{\text{١ + ٠.٦٧}} = \text{٠.٦٠}$$

وهنا يمكن أن نجد القيم المطلوبة في الصيغة العامة للمعادلة الخاصة بتقدير القدرة ومنها احتمال الإجابة الخاطئة(ح ع) حيث أنها متممة احتمال الإجابة الصحيحة وكذلك القيمة ت س (ن م − ح س(ق ن)) والقيمة ت² ر ح (س) ح ع (ع) وذلك على النحو التالي

ح ع = ١ − ح س = ١ − ٠٫٦٠ = ٠٫٤٠

ت س (ن م − ح س(ق ن)) = ٢ ١(صفر − ٠٫٦٠) ٢ ١ (−٠٫٦٠) = −٠٫٧١٥٢

ت² ر ح (س) ح ع (ع) = ١.٢ ² (٠٫٦٠)(٠٫٤٠) = ١٫٤٤ (٠٫٢٤) = ٠٫٣٤٦٧

ثالثا: بالنسبة للفقرة الثالثة وعند مستوى القدرة ق =٠.٣٢٤، ص= ١، ت = ٠٫٨ ن م = ١ فان إجراءات التدوير الثالث تتم على النحو التالي

ح ص١ = ١ / (١ + ت − ت(ق − ص)) =

ح ص٣ = ١ / (١ + ٢٫٧١٨ − ٠٫٨ (٠٫٣٢٤ − ١))

ح ص٣ = ١ / (١ + ٢٫٧١٨ (٠٫٥٤٠٨))

ح ص٣ = ١ / (١ + ١٫٧١) = ٠.٣٧

وهنا يمكن أن نجد القيم المطلوبة في الصيغة العامة للمعادلة الخاصة بتقدير القدرة ومنها احتمال الإجابة الخاطئة ح ع حيث أنها متممة احتمال الإجابة الصحيحة وكذلك القيمة ت س (ن م − ح س(ق ن)) والقيمة ت² ر ح (س) ح ع (ع) وذلك على النحو التالي

ح ع = ١ − ح س = ١ − ٠٫٣٧ = ٠.٦٣

ت س (ن م – ح ص(ق ر)) = ٨٫٠ (١ – ٠٫٣٧) (٠٫٦٣) = ٠٫٥٠٥٦

ت²ر ح (ص) ح خ (خ) = ٨٫٠ ²(٠٫٣٧)(٠٫٦٣) = ٦٤٫٠ (٠٫١٣٣٢) = ٠٫١٤٨٨

والآن يمكن تلخيص البيانات التي حصلنا عليها كما في الجدول رقم (١١) أدناه لتقدير قدرة المفحوصين ولكن باستخدام معلمة القدرة المقدرة (ق = ٠٫٣٢٤) نطبق الصيغة العامة لتقدير معلمة القدرة.

جدول (١١) : القيم الخاصة بتقدير قدرة المفحوصين في الفترة الثالثة للمثال (١)

الفترة	ن م	ح ص	ح م	ت(ن م – ح ص)	ت²(ح ص * ح م)
١	١	٠٫٧٩	٠٫٢١	٠٫٢١٠٢	٠٫١٦٦
٢	٠	٠٫٦٠	٠٫٤٠	٠٫٧١٥٢-	٠٫٣٤٦٧
٣	١	٠٫٣٧	٠٫٦٣	٠٫٥٠٥٦	٠٫١٤٨٨
المجموع				٠٫٠٠٠٦	٠٫٦٦١٥

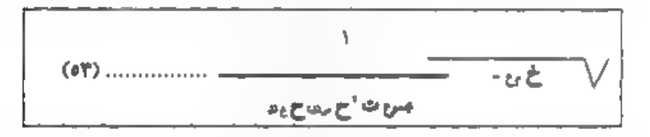
وكما تم في عملية الفلترة الثانية سنقوم بإعادة نفس الإجراءات لتقرير الاستمرار أو التوقف بالنسبة لعملية الفلترة ولكن مع الانتباه أيضا إلى قيمة معلمة القدرة المستخدمة وهي القيمة (ق = ٠٫٣٢٤) وذلك على النحو التالي:

$$ق_{ر+1} = ق_{ر} + \frac{\text{مجموع } - \text{ت س (ن م – ح ص(ق ر))}}{\text{مجموع ت٢ر (ح ص (ق ر)) (ح خ (ق ر))}}$$

$$ق_{ر+1} = ٠٫٣٢٤ + \frac{(٠٫٠٠٠٦)}{(٠٫٦٦١٥)} = ٠٫٣٢٤ + (٠٫٠٠٠٩) = ٠٫٣٢٤٩$$

من الملاحظ إن قيمة معلمة القدرة للمفحوصين هنا زادت عن القيمة في الفلترة الثانية، لكن ما هو حجم هذه الزيادة والتي يمكن إيجادها من خلال إيجاد الفرق بين المقدارين في الفلترة الثانية والثالثة أي أن (٠٫٣٢٤٩ – ٠٫٣٢٤٠ = ٠٫٠٠٠٩) وهو

مقدار بسيط جدا أي أن القيمة المضافة نتيجة لعملية الفلترة الثالثة لم تزيد عن القيمة (٠.٠٠٠٩) وهي زيادة غير مجدية أو غير عملية وهذا يشير إلى أو يؤكد بأن القدرة الحقيقية للمفحوص لا يمكن التوصل إليها بشكل مطلق والقيمة التي سنحصل عليها ستكون تقديرا لها، وبالنسبة لمثالنا فان القدرة التي قمنا بتقديرها تمثل أفضل تقدير لقدرة المفحوصين نتيجة لخضوع المفحوصين للاختبار لعدة مرات مع الافتراض بالطبع بعدم وجود أثر لعامل التذكر للإجابات على الفقرات نتيجة لمرات التطبيق السابقة، وبما أنه سيتم تقدير القدرة الحقيقية لهم عند كل مرة تطبيق للاختبار فانه يمكن التأكد من مدى دقة هذا التقدير من خلال إيجاد قيمة الخطأ المعياري للتقدير (Standard Error of Estimating) وذلك من خلال المعادلة (٥٣)



من الملاحظ أن القيمة التي تحت الجذر الذي يمثل مقام المعادلة رقم (٥ ـ ٢) تمثل مقام معادلة الصيغة العامة لتقدير معلمة القدرة للمفحوصين (٥ ـ ١) حيث تتضمن كل من معلمة التميز واحتمال الإجابة الصحيحة على الفقرة ومتممتها، ويمكن من خلال المعادلة (٥-٢) الحكم على مدى دقة التقدير لمعلمة قدرة المفحوصين في كل مرة فلترة أو لآخر مرة فلترة نريد التوقف عندها، إ ذا أنه كلما زادت هذه القيمة كلما كان ذلك مؤشرا على انخفاض دقة التقدير، ففي المثال السابق لو طبقنا المعادلة رقم (٥ ـ٢) لوجدنا أن

$$خ_{ع} = \frac{1}{٠.٦٦١٥} = ١.٢٢٣$$

وتعتبر هذه القيمة عالية حيث تعني أن دقة التقدير متدنية لان قيمة خطأ التقدير المعياري للتقدير عالية، إضافة إلى أن تدني قيمة الخطأ المعياري للتقدير قد تتأثر بعدد

الفقرات التي يتكون منها الاختبار، علما أن إجراءات تقدير معلمة القدرة للمفحوصين قد تفشل في الحصول على تقدير دقيق للقدرة وذلك في حالتين .

الأولى: عندما يفشل المفحوصين في الإجابة على أي من الفقرات التي يتكون منها الاختبار، ويزداد الأمر صعوبة كلما زاد عدد الفقرات غير المجابة من جميع المفحوصين، ويحدث ذلك عندما تكون هذه الفقرات صعبة

الثانية: عندما يجيب جميع المفحوصين على أي من الفقرات ويزداد الأمر صعوبة كذلك كلما زاد عدد الفقرات المجاب عليها من قبل جميع المفحوصين ، ويحدث ذلك عندما تكون الفقرات سهلة .ويكمن سبب عدم إمكانية التقدير الدقيق لمعلمة القدرة في الحالتين السابقتين بسبب الاستبعاد الذي يتم للفقرات الصعبة كما في الحالة الأولى، وكذلك الفقرات السهلة كما في الحالة الثانية، حيث لا تدخل في التحليل وبالتالي في تقدير المعلمة، حيث تكون برامج التحليل الإحصائي مهيأة بحيث لا تدخل الفقرات التي يجيب عليها جميع المفحوصين (السهلة) أو تلك التي لم يجب عليها أي من المفحوصين (الصعبة) في عملية التحليل

علاقة ثبات الفقرة بتقدير قدرة المفحوصين

من الأمور المهمة في نظرية السمات الكامنة هو ثبات القدرة من خلال الفقرات التي تتعلق بها ويعتمد ثبات الفقرة على عاملين هامين

- قياس جميع فقرات الاختبار لنفس السمة
- وقوع جميع معالم الفقرات على وحدة قياس واحدة (متصل موحد) .

ولتوضيح ذلك افترض أن علامة مفحوص ما (صفر) والتي تضعه على منتصف متصل السمة، وتم تطبيق مجموعة من الفقرات التي تتمتع بمعامل صعوبة (- ٢) فانه يمكن استخدام هذه الفقرات في تقدير قدرة المفحوص، وكذلك تم تطبيق مجموعة أخرى من الفقرات التي تتمتع بمعامل صعوبة (١) فان هذه الفقرات يمكن استخدامها في تقدير قدرة المفحوص، بمعنى ان هناك اختبارين وكذلك هناك مستويين من القدرة،

فكأن هناك مجموعتين ولكل مجموعة مستوى قدرة ق ١ ق ٢، وفي ظل مبدأ الثبات فانه يمكن القول أن ق ١ = ق ٢ وهذا يعني أن كلا الاختبارين يجب أن يؤديا إلى تقديرات متقاربة لقدرة المفحوصين اللذين خضعوا لكلا الاختبارين حيث تتمتع الفقرات فيهما بمستوى صعوبة مختلفين، كما أنه ليس من الضروري أن تتمتع الفقرات بنفس المستوى من التمييز، وهذا يعكس الحقيقة التي تقول "ان منحنى خصائص الفقرة هو انعكاس لمقياس القدرة" لذلك فان اي جزء من متصل القدرة يمكن أن يدخل في تقدير معلمات الفقرة، حيث أن هناك عدة مقاطع من منحنيات خصائص الفقرات التي تمثل أو يمكن استخدامها في تقدير قدرات المفحوصين.

فالفقرات ذات الصعوبة العالية ستمثل نقطة على منحنيات خصائصها والتي تتطابق مع القدرة المناسبة لها، وكذلك الفقرات ذات الصعوبة المتدنية فستمثل نقطة على منحنيات خصائصها، وكلا النوعين من الفقرات يمكن استخدامها في تقدير قدرات المفحوصين عند كل نقطة يمثلها كل نوع من الفقرات من حيث صعوبتها، وفي كلا الحالتين فان منحنى خصائص الفقرة يجب إن يتوفر لكن هذا من غير المنطقي أن يتم.

إن التطبيق العملي لهذه الحقيقة او المبدأ هو أن أي اختبار يمكن تحديده أو تعينه على متصل السمة (القدرة) يمكن استخدامه في تقدير قدرة المفحوصين، فعلى سبيل المثال اذا ما طبق اختبار سهل أو صعب على مجموعة من المفحوصين فيمكن من خلالهما تقدير القدرة، وهذا ما يغاير مبدأ النظرية الكلاسيكية حيث يحصل المفحوصين على علامة عالية في الاختبار السهل وعلى علامة متدنية في الاختبار الصعب، وهذا ما يعيق الحصول على القدرة الحقيقية للمفحوصين وبشكل دقيق ، أما في النظرية الحديثة فان قدرة المفحوص ثابتة بغض النظر عن الفقرات التي يخضعون لها، ومعنى كلمة دقيق أن القدرة لا تتغير بتغير صعوبة الفقرات أو سهولتها، وعلى سبيل المثال فانه إذا خضع المفحوصين للاختبار أكثر من مرة وعلى افتراض أنه لا اثر لعامل التذكر من مرة تطبيق لمرة تطبيق لاحقة فان القدرة ستبقى ثابتة، أما إذا تعرض المفحوصين إلى مواقف تعليمية بين التطبيق الأول والتطبيق والآخر أو تأثروا بكل تطبيق سابق فان أداء أو قدرة المفحوصين ستتغير من تطبيق إلى آخر وبالتالي ستختلف قدراتهم من

دار جرير للنشر والتوزيع

تطبيق إلى آخر وعلى نفس الاختبار. ولذلك فان القدرة الحقيقية للمفحوصين حسب تطبيقات النظرية الحديثة تبقى ثابتة حتى لو تغيرت الفقرات من حيث الصعوبة والسهولة، ولذلك فان هناك تطبيقات عملية دقيقة ومهمة لهذه النظرية ومنها ثبات قدرة المفحوصين أو ثبات المجموعات على الفقرات مهما اختلفت صعوبة فقرات الاختبارات.

خلاصة

من خلال الفصل السابق هناك مجموعة من الاعتبارات المهمة تم تصنيفها على النحو التالي.

أولا: توزيع القدرة المقدرة.

- إن المعدل المنطقي للتقديرات يجب أن يقترب من معلمة القدرة للمفحوصين من خلال البرامج الحاسوبية.
- إذا كانت صعوبة الفقرات تقترب من قدرات المفحوصين فان متوسط تقديرات القدرات للمفحوصين يجب أن يقترب من قيمة القدرة
- تكون قيمة الخطأ المعيار ي للتقدير عالية إذا كانت قيم صعوبة الفقرات غير قريبة من قيم قدرات المفحوصين، وفي هذه الحالة فان القيم النظرية للأخطاء المعيارية للتقديرات ستكون عالية كما ستقترب القيم المحسوبة من القيم الملاحظة
- إذا كانت قيم معاملات التميز للفقرات عالية فان قيمة الخطأ المعياري للتقدير ستكون متدنية وعلى العكس إذا كانت قيم معاملات التميز للفقرات متدنية فان قيمة الخطأ المعياري ستكون عالية .
- إن الوضع المناسب لتقدير قدرة المفحوصين هو أن تكون قيم معلمات الصعوبة للفقرات متقاربة من جهة، وتتساوى كل من صعوبة الفقرات مع معلمة القدرة وأما قيم معلمات التميز فيجب أن تكون عالية .

ثانيا: علاقة ثبات الفقرة بتقدير قدرة المفحوصين.

- إن الفقرات المختلفة من حيث صعوبتها تؤدي إلى قيم مقدرة للقدرة بحيث تقترب من مستوى القدرة الحقيقية للمفحوصين.



- يميل نمط التقديرات بشكل عام إلى الاقتراب من معلمة قدرة المفحوصين، وإذا طبقت مجموعة من الاختبارات، بحيث يكون عدد الفقرات فيها كبيرا، حيث ينزع متوسط تقديرات القدرة إلى أن يتساوى مع معلمة قدرة المفحوصين، وهذا يؤكد ان طول الاختبار يزيد من ثباته (عوده، ٢٠٠٥) إضافة إلى أنه في التقديرات المتقاربة والتي قد تظهر بأنها تشكل عنقودا واحدا يتمحور حول قيمة معلمة القدرة، وفي بعض الحالات يبدو ان ثبات الفقرة قد ينتهك أو يتم تجاهله.

ثالثا: ثبات قدرة المفحوصين .

- يؤدي اختلاف معلمات الفقرات إلى قيم مختلفة لتقديرات القدرة وبالرغم من ذلك فان هذه التقديرات، لا بد أن تقترب من القيمة الحقيقية للقدرة.
- تميل قيم التقديرات إلى الاقتراب من معلمة القدرة، وإذا استخدمت عدة اختبارات تتكون من عدد كبير من الفقرات (اختبارات طويلة) فان متوسط تقديرات القدرة سيكون مساويا لمعلمة قدرة المفحوصين، كما تميل هذه التقديرات إلى التجمع في مجموعات حول قيمة المعلمة وفي هذه الحالة فان مبدأ ثبات الفقرة يكون قد تم انتهاكه.
- من الملاحظ أن الأمثلة التي تم استخدامها سابقا لتوضيح الإجراءات الرياضية المختلفة قد أبرزت مفهومين يتعلق الأول بتغير تقديرات القدرة حول معلمة القدرة للمفحوصين، وهذا وضح كيفية تقدير الاختبار لقدرة المفحوصين على متصل السمة، أما المفهوم الثاني فيتعلق بثبات الفقرة في تقدير قدرة المفحوصين وهو تقديم للفصل التالي للفصل القادم ، ومن المهم ان نشير إلى أنه على القارئ أن يأخذ بعين الاعتبار أن تقدير القدرة هو شكل آخر من علامات الاختبار لكنها تفسر في إطار النظرية الحديثة للقياس.
- في الفصل الأول تم الحديث عن مفهوم السمة الكامنة، ومن الإجراءات المكملة لهذه النظرية أنه يمكن تحديد مواقع المفحوصين على متصل السمة، ولذلك فان لكل مفحوص علامة قدرة (قيمة المعلمة) تحدد موقعه على متصل السمة. وعلى

الرغم من ذلك فانه لا يمكن الحصول على قيمة معلمة القدرة بشكل مطلق، وإننا في أفضل الأحوال نلجأ إلى تقديرها ومن خلال الإجراءات الحاسوبية يمكننا اشتقاق تقدير معلمة القدرة للمفحوصين، حيث أن الإجراءات الحاسوبية تتم بناء على افتراض إمكانية اشتقاق هذا التقدير، لكن هذا الافتراض لا يمكن البرنامج الحاسوبي من اشتقاق العوامل الخاصة باستجابة الفقرة التي يمكن من خلالها الحصول على تقديرات القدرة وتوضيح بنية نظرية السمات الكامنة

الفصل السادس

اقتران المعلومات

Information Function

يقصد بمفهوم المعلومات الوارد في عنوان هذا الفصل المعلومات أو المعرفة التي يمكن الحصول عليها إما عن المفحوصين أو عن الفقرات من خلال تطبيق الاختبار مثل معلمات الفقرة وقدرة المفحوصين سواء أحصلنا عليها من خلال منحنى خصائص الفقرة أو منحنى خصائص الاختبار أي أنها المعلومات التي تتوفر من الفقرة أو الاختبار عن المفحوص، ومن الجدير بالذكر أن فيشر (Fisher) هو أول من طرح مفهوم المعلومات بهذا المعنى، حيث عرّف المعلومات على أنها مدى دقة المعلومات التي يمكن معرفتها أي أنه لا بد من الدقة كمعيار للمعلومات التي يمكن الحصول عليها عن الفقرات أو الاختبار أو عن المفحوصين، ويمكن قياس الدقة من التغير في التقديرات التي تتعلق بمعلمة محددة، ولذلك فان قياس الدقة يتم من خلال فكرة تباين التقديرات والذي سنرمز له بالرمز ($ع^2$) ويقارن مقدار الدقة بالقيمة (١)، ويمكن الحكم على دقة المعلومات من خلال المعادلة (٥٤) .

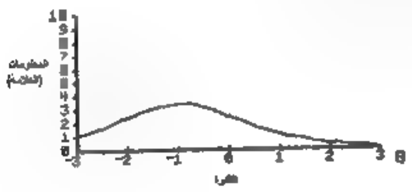
$$د = \frac{١}{ع^2} \quad \text{.........................} (٥٤)$$

حيث :

د : دقة المعلومات

$ع^2$: تباين تقديرات المعلمات.

وينصب الاهتمام حسب نظرية السمات الكامنة على تقدير معلمة القدرة بالنسبة للمفحوص، ، وكما هو معروف فان معلمة القدرة يرمز لها (ق) وان تقدير معلمة القدرة يرمز له بالرمز(قˆ)أي أن (قˆ) هو تقدير ل(ق)، حيث يمكن إيجاد الانحراف المعياري لتقديرات معلمات القدرة للمفحوصين، وإذا تم تربيع هذه القيمة فسنحصل على تباين هذه التقديرات، وهو مؤشر على مستوى الدقة في تقدير مستوى القدرة، ويعبر عن مدى دقة المعلومات التي يمكن الحصول عليها من عملية التقدير من خلال مفهوم التبديل لتباين التقديرات والتبديل هو (١ / التباين)، فإذا كانت القيمة كبيرة فان المفحوص الذي لديه قدرة حقيقية في مستوى قدرة ما يمكن تقديرها بشكل دقيق، وبذلك فان معظم التقديرات التي يمكن الحصول عليها تكون قريبة من القدرة الحقيقية، وعلى العكس من ذلك فإذا كانت القيمة متدنية فان ذلك مؤشر على عدم إمكانية تقدير القدرة بشكل دقيق بمعنى أن التقديرات ستكون بعيدة عن القدرة الحقيقية، ويمكن استخدام المعادلة (٦ ــ ١) في حساب قيمة أو مقدار المعلومات عند كل مستوى قدرة على متصل القدرة ضمن المدى (- ∞ ــ ∞)، ولان القدرة هي متغير متصل فان مقدار المعلومات سيكون متغيرا متصلا وإذا تم رسم مقدار أو شكل المعلومات مقابل القدرة فان الشكل سيكون كما هو في شكل (٣١) حيث يتبين منه مقدار المعلومات التي تقدمه فقرة ما.



شكل (٣١) منحنى اقتران المعلومات للفقرة

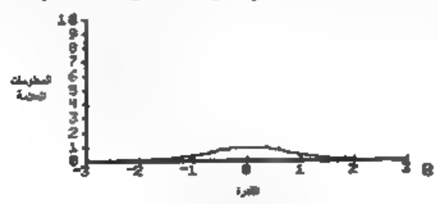
يتضح من الشكل أعلاه إن الفقرة تقدم أعلى قدر من المعلومات عن المفحوصين من ذوي مستوى القدرة (- ١) حيث تبلغ القيمة الموازية لأعلى قدر من المعلومات عند قمة منحنى (اقتران) المعلومات وتبلغ قيمة المعلومات (٣)، ويمكن معرفة ذلك من خلال مد خط عمودي من أعلى قمة للمنحنى ليصل إلى متصل القدرة (محور السينات) وخط أفقي من نفس النقطة (أعلى قمة للمنحنى) نحو متصل المعلومات (محور الصادات) فتكون القيمة التي تتقاطع مع الخط الأفقي تمثل أعلى قيمة للمعلومات التي تقدمها الفقرة عن المفحوصين وذلك عند مستوى القدرة الذي يتقاطع مع الخط العامودي، ومن خلال الشكل تزداد قيمة المعلومات كلما انتقلنا من مستوى القدرة (-٢) نحو مستوى القدرة (-١)، ومن الملاحظ أن مقدار المعلومات يقل بشكل تدريجي والنقطة التي ينخفض المنحنى عندها المنحنى بشكل سريع ومفاجئ تقل إمكانية التقدير ولن تكون بالشكل المرغوب أو المطلوب، وبمعنى آخر فان الشكل أعلاه يبين لنا ما هي الدقة التي يمكننا من خلالها تقدير المعلمات عند كل مستوى من مستويات القدرة، وذلك اعتمادا على المعلومات التي تقدمها الفقرة حيث تقل هذه المعلومات كلما ابتعدنا عن مستوى القدرة (-١) وخاصة إذا انتقلنا إلى مستويات القدرة الأعلى من (-١) .

ومن الجدير بالذكر أن اقتران معلومات الفقرة لا يعتمد على توزيع المفحوصين على متصل القدرة وهو ما يعتمد عليه منحنى خصائص الفقرة ومنحنى خصائص الاختبار، وبشكل عام فان قيمة اقتران المعلومات تبرز وتتضح من خلال مد خط افقي من أي نقطة على المنحنى بحيث تقابل مستوى قدرة معين على متصل المعلومات حيث سيقابل القدرة التي تقدمها الفقرة ذلك المستوى من القدرة ، ففي الشكل (٣١) فان النقطة (جـ) مثلا تمثل أعلى قمة لمنحنى اقتران المعلومات وهذا يعني أن أعلى مدى دقة تقدمها هذه الفقرة ستكون عن المفحوصين من ذوي مستوى القدرة (- ١) وهو مستوى القدرة الذي تقابله أعلى نقطة للمنحنى النقطة (جـ) وهي قمة المنحنى وقد بلغ مقدار المعلومات التي تقدمها هذه الفقرة (٣) وتزداد صعوبة الحصول على معلومات دقيقة عن المفحوصين كلما ابتعدنا عن مستوى القدرة (- ١) نظرا لانخفاض مقدار المعلومات الذي تقدمه بسبب انخفاض منحنى اقتران المعلومات الذي

يعني انخفاض مقدار المعلومات عند تلك المستويات، ومن المهم أن يدرك القارئ أن الوعي بالأفكار والمعلومات التي وردت سابقا ضروري ومهم جدا لكل من باني الاختبار ومستخدمه على حد سواء، لان ذلك يعني الدقة التي تتمتع بها عملية التقدير لقدرة المفحوصين والتي تعتمد على موقع المفحوصين من متصل القدرة

اقتران معلومات الفقرة Item Information Function

بما أن منحنى اقتران المعلومات يعتمد على الفقرات التي يتكون منها الاختبار سواء كانت منفردة (منحنى خصائص الفقرة) أو مجتمعة (منحنى خصائص الاختبار)، فان النظرية الحديثة تسمى أحيانا بنظرية الفقرة، ووفقا لهذه النظرية فان كل فقرة تقيس سمة محددة ولذلك فان مقدار المعلومات يعتمد على الفقرة الواحدة ويمكن حساب هذا المقدار (المعلومات) عند أي مستوى من خلال ما يمكن الإشارة إليه ب م ف ه، حيث ترمز (م) إلى المعلومات ف إلى الفقرة ولان الفقرة هي فقرة واحدة فان مقدار المعلومات عند أي نقطة (مستوى) من متصل القدرة سيكون منخفض، لأنه في الواقع سيكون جزءا من منحنى اقتران المعلومات للاختبار عند مستوى قدرة معين، وفيما لو تم رسم مقدار المعلومات سيتخذ نمطا كما في الشكل (٦ ــ ٢). حيث تبلغ قيمة مقدار المعلومات الذي تقدمه الفقرة أعلى حد لها عند مستوى القدرة الذي يساوي صعوبة الفقرة، وتقل هذه القيمة كلما ابتعدنا عن ذلك المستوى وهو بالنسبة للشكل (٣٢) المستوى (صفر) حيث يقابل أعلى قمة لمنحنى اقتران المعلومات للفقرة.



الشكل (٣٢) منحنى اقتران المعلومات لفقرة ما

اقتران معلومات الاختبار: Test Information Function

بالرغم من أن الاختبار بشكل عام يستخدم لتقدير قدرة المفحوصين، إلا أن مقدار المعلومات الذي يمكن توفيره أو الحصول عليه يمكن ان يتم تحديده عند أي مستوى من مستويات القدرة، وبما أن الاختبار هو مجموعة من الفقرات، فان معلومات الاختبار عند مستوى معين من مستويات القدرة هي مجموع معلومات الفقرات عند ذلك المستوى من القدرة وبذلك يمكن تعريف اقتران معلومات الاختبار من خلال المعادلة (٥٥) .

م (ق) = مجموع م ل (ق) ....(٥٥)

حيث

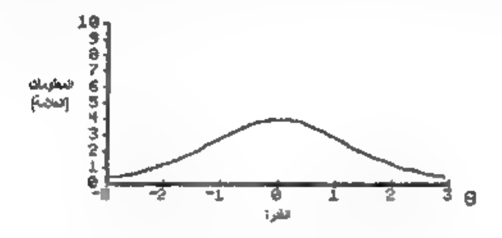
م (ق) :مقدار معلومات الاختبار عند مستوى القدرة (ق)

م ل (ق) : مقدار معلومات الفقرة عند مستوى قدرة (ق).

ن : عدد فقرات الاختبار.

وبما أن الاختبار مكون من مجموعة من الفقرات ومن خلال المعادلة أعلاه فان مقدار المعلومات الخاص بالاختبار يجب أن يكون أعلى من مقدار المعلومات الخاص بالفقرة الواحدة، ومن هنا فان الاختبار يقيس القدرة أكثر من مما تقيسه الفقرة، ومما هو جدير بالاهتمام في تعريف مفهوم اقتران المعلومات الوارد في المعادلة السابقة أنه كلما زاد عدد فقرات الاختبار زاد مقدار المعلومات التي يمكن الحصول عليها عن المفحوصين، ومن هنا فالاختبار الطويل (عدد الفقرات كبير) يعطي معلومات أكثر من الاختبار القصير (عدد الفقرات قليل) حيث يتضمن الشكل (٣٣) منحنى اقتران المعلومات لاختبار يتكون من (١٠) فقرات حيث يبين الشكل أن مقدار الملومات التي تم الحصول عليها تعتبر قيمة متوسطة، وفيه يتناقص مقدار المعلومات كلما ابتعدنا عن مستوى القدرة الذي يعطي فيه الاختبار أكبر قدر من المعلومات أي عند المستوى القدرة (ق = صفر)، ولذلك فان التقدير يكون أكثر دقة في تلك النقطة (ق= صفر)

وهي نقطة منتصف متصل السمة لاحظ قمة المنحنى، وهذا يعني أن مقدار المعلومات وبالتالي دقة تقدير القدرة يتناقص كلما اقتربنا من مستويات القدرة العالية والمتدنية.



شكل (٣٣) منحنى اقتران الاختبار

ويعتبر اقتران معلومات الفقرة من التطبيقات العملية والمفيدة للنظرية الحديثة في القياس، حيث يبين كيف يسير الاختبار من حيث تقدير القدرة من خلال علامات القدرة، وعندما يكون اقتران معلومات الاختبار بمثابة الخط الأفقي فان ذلك يعني أن الاختبار يهدف إلى غرض محدد،، فعلى سبيل المثال عندما يستخدم الاختبار لترشيح الطلبة لمنحة تعليمية مثلا فان الوضع النموذجي لذلك الاختبار قد لا يتحقق، وفي هذه الحالة فانه لا بد من قياس القدرة مع الأخذ بعين الاختبار مفهوم الدقة لمستويات القدرة عند الحد الذي يحدد من يحصل أو لا يحصل على المنحة(درجة القطع)، وفي هذه الحالة فان أفضل اقتران لمعلومات الاختبار ستكون قمته مقابل نقطة القطع، وعندما نريد استخدام الاختبار لأغراض خاصة أخرى فإننا بحاجة إلى نماذج أخرى من اقتران معلومات الاختبار، أي أن الغرض من الاختبار يحدد أفضل اقتران لمعلومات الاختبار.

وكما هو معروف فانه يمكن إيجاد اقتران المعلومات لكل فقرة في الاختبار كل على حدة، وان مقدار المعلومات يمكن الحصول عليه من كل فقرة سيكون محدودا

(قليلا) إذا ما قورن بمقدار المعلومات التي يمكن الحصول عليه من الاختبار. إضافة إلى أننا لا نحاول تقدير القدرة للمفحوصين من فقرة واحدة، وهكذا فان إيجاد مقدار معلومات الاختبار عند مستوى قدرة معين يتم الحصول عليه من خلال جمع معلومات الفقرات عند كل مستوى قدرة ولذلك فان التعريف الرياضي لمقدار هذه المعلومات يعتمد على النموذج اللوغريتمي المستخدم لمنحنى خصائص الفقرة، ولذلك من الضروري أن يتم اختيار مقدار المعلومات عند كل النماذج لتحديد أفضل النماذج التي يمكن أن يوفر الاختبار من خلالها أعلى مقدار من المعلومات

تعريف اقتران معلومات الفقرة.

إن مقدار المعلومات التي يمكن الحصول عليها من الاختبار يتأثر بالنموذج المستخدم، حيث أن لكل نموذج افتراضاته ومعلماته ومن هنا فان تعريف معلومات الفقرة يختلف باختلاف النموذج المستخدم، وهذا يعني أن تعريف وحساب معلومات الفقرة يختلف من نموذج إلى آخر، وتبعا لذلك تختلف طريقة حسابه من نموذج إلى آخر حيث سنورد هذه التعريفات وفقا للنماذج الثلاثة على النحو التالي.

اولا: النموذج احادي المعلمة.

في ظل النموذج آحادي المعلمة فان التعريف الرياضي لاقتران معلومات الفقرة يكون حسب المعادلة (٥٥)

م ف(ق) = ح ص ف(ق) ح خ ف(ق) .... (٥٥)

حيث :

ح ص ف(ق) : احتمال الإجابة الصحيحة على الفقرة.

ح خ ف(ق) : احتمال الإجابة الخاطئة على الفقرة.

ولتوضيح كيفية تطبيق المعادلة (٦ ــ ٣) سوف يتم حل المثال التالي

دار جرير للنشر والتوزيع

مثال (١).

لنحسب اقتران معلومات الفقرة عندما تكون قيمة معلمة التمييز لها (ث = ١) وقيمة معلمة الصعوبة لها (ق = ١) وذلك عند مستويات القدرة من (-٣ ـ ٣)؟

الحل :

إن حساب اقتران معلومات الفقرة يتطلب حساب القيم التالية

أولا: احتمال الإجابة الصحيحة على الفقرة.

$$\text{ح ف(ق)} = \frac{١}{١ + \text{ث} - ١(-٣ - ١)}$$

$$\text{ح ف(ق)} = \frac{١}{١ + \text{ث} - ١(- ٤)}$$

$$\text{ح ف(ق)} = \frac{١}{١.+ ٢.٧١٨ ٤}$$

$$\text{ح ف(ق)} = \frac{١}{١ + ٥٤.٥٧}$$

$$\text{ح ف(ق)} = \frac{١}{٥٥.٥٧} = ٠.٠١٧٠ \sim ٠.٠٢$$

ثانيا . حساب قيمة احتمال الإجابة الخاطئة من خلال المعادلة التالية

$$\text{ح خ ف(ق)} = ١ - \text{ح ص ف(ق)}$$

$$= ١ - ٠.٠٢ \quad = ٠.٩٨$$

ثالثا: حساب قيمة اقتران معلومات الفقرة من خلال المعادلة (٦ ـ ٣) على النحو التالي:



م (ق) = (ح ص ف (ق)) ● (ح ص ف (ق)) .

= (٠.٠٢) ● (٠.٩٨)

= ٠.٠٢

وعند تطبيق نفس الإجراءات على مستويات القدرة الأخرى فان اقتران معلومات الفقرة والقيم اللازمة لحسابه ستكون كما في الجدول رقم (١٢) على النحو التالي .

جدول(١٢) : قيم المعلمات والقيم اللازمة لحساب اقتران معلومات الفقرة في المثال (١).

ق	ـ ل	١ + هـ ^ل	ح ص	ح ع	ح ص ● ع	ت ² (*)	م(ق)
-٣	-٤	٥٥.٥٧	٠.٠٢	٠.٩٨	٠.٠٢	١	٠.٠٢
-٢	-٣	٢٠.٠٩	٠.٠٥	٠.٩٥	٠.٠٥	١	٠.٠٥
-١	-٢	٧.٣٩	٠.١٢	٠.٨٨	٠.١١	١	٠.١١
صفر	-١	٢.٧٢	٠.٢٧	٠.٧٣	٠.٢٠	١	٠.٢٠
١	صفر	١	٠.٥٠	٠.٥٠	٠.٢٥	١	٠.٢٥
٢	١	٠.٣٧	٠.٧٣	٠.٢٧	٠.٢٠	١	٠.٢٠
٣	٢	٠.١٤	٠.٨٨	٠.١٢	٠.١١	١	٠.١١

من الملاحظ أن قيمة (اقتران) المعلومات الناتجة لمثل هذه الفقرة أقل منه بالنسبة للفقرة في المثال رقم (١)، وهذا انعكاس لقيمة معلمة التمييز للفقرة حيث ان قيمة معلمة تمييز الفقرة في هذا المثال (ت = ١) أقل منه للفقرة في المثال (١) والذي بلغت قيمته (ت = ١.٥) ولذلك فان قيمة أو مقدار المعلومات ستكون أقل لان الفقرة ذات التمييز المتدني ستقدم معلومات اقل من تلك التي يكون تمييزها عالي، ومن المهم أن ندرك أن مقدار المعلومات الذي تقدمه الفقرة سوف يقع حول قيمة مستوى القدرة الذي يساوي صعوبتها، ومن الملاحظ أنه عندما تكون قيمة معلمة التمييز (ت=١) عند

* في النموذج أحادي المعلمة تكون قيمة معلمة التمييز ثابتة وقيمتها (١) لجميع الفقرات.

مستوى معين، فان مقدار المعلومات لكل فقرة يساوي حاصل ضرب احتمالي الإجابة الصحيحة والخاطئة عند ذلك المستوى، وهـذا هـو السبب في صيغة معادلة اقتران المعلومات في النموذج أحادي المعلمة .

ثانيا : النموذج ثنائي المعلمات.

يعرف اقتران معلومات الفقرة في ظل النموذج ثنائي المعلمات حسب المعادلة (٥٧) على النحو التالي:

م ف(ق) = ت^٢ ح ص ف(ق) ح ع ف(ق) ....(٥٧)

حيث

م ف(ق) : اقتران معلومات الفقرة عند مستوى قدرة معين.

ت: معامل تمييز الفقرة.

ح ص ف(ق) احتمال الإجابة الصحيحة على الفقرة عند مستوى قدرة معين.

ويحسب من خلال المعادلة (٢ ــ ٢) ح (ق) = ١/١ + ث ^(-ت(ق - ص))

ح ع ف(ق): احتمال الإجابة الخاطئة على الفقرة عند مستوى قدرة معين.

ويحسب من خلال المعادلة ح ع (ق) = (١− ح ص(ق))

ق : مستوى القدرة قيد الاهتمام.

ولتوضيح كيفية حساب اقتران معلومات الفقرة عند مستويات القـدرة المختلفة سوف نقوم بحل المثال التالي .

مثال (٢).

اوجد اقتران معلومات الفقرة التي تتمتع بمعامل صعوبة (ص =١) ومعامل تمييز (ت = ١٫٥) عند مستويات القدرة من (− ٣ ــ ٣)

الحل:

سوف نقوم بالحل بحيث نحسب القيم اللازمة عند مستوى قدرة (– ٣) وتطبيق الإجراءات نفسها على المستويات الأخرى كما تظهر في الجدول (١٠) وذلك على النحو التالي:

أولاً. نحسب احتمال الإجابة الصحيحة على الفقرة

$$\text{ح ف(ق)} = \frac{١}{١ + \text{ث}^{-\text{ت(ق} - \text{س)}}}$$

$$\text{ح ف(ق)} = \frac{١}{١ + ٢.٧١٨^{-١.٥(-٣-١)}}$$

$$\text{ح ف(ق)} = \frac{١}{١ + ٢.٧١٨^{٦}}$$

$$\text{ح ف(ق)} = \frac{١}{١ + ٤٠٣.١٧}$$

$$\text{ح ف(ق)} = \frac{١}{٤٠٤.١٧} = ٠.٠٠٢٤ \sim \text{صفر}$$

ثانيا: سنقوم بحساب مقدار المعلومات عند هذا المستوى من القدرة من المعادلة (٦-٤).

$$\text{م ن(ق)} = \text{ت}^{٢}\ \text{ح ص ف(ق)}\ \text{ح خ ف(ق)}$$

وهذا يتطلب حساب القيم التالية

– احتمال الإجابة الخاطئة من خلال المعادلة التالية

$$\text{ح خ ف(ق)} = ١ - \text{ح ص ف(ق)}$$

$$= ١ - \text{صفر} = ١$$

– حاصل ضرب احتمال الإجابة الصحيحة واحتمال الإجابة الخاطئة على الفقرة

(ح س ف(ق)) * (ح ع ف(ق))

١ * صفر = صفر

- مقدار المعلومات عند (ق = -٣)

م ف(ق) = ت^٢ ح س ف(ق) ح ع ف(ق)

=(١.٥) ^٢ *(صفر) (١)

= صفر

وهنا سنكرر العملية عند جميع مستويات القدرة الأخرى وستظهر القيم كما في الجدول رقم (١٣) على النحو التالي.

جدول(١٣): قيم المعلمات والقيم اللازمة لحساب اقتران معلومات الفقرة في المثال (٢).

ق	- ل	١ + هـ^-ل	ح س	ح م	ح س * ح م	ت^٢	م(ق)
-٣	-٦	٤٠٣.١٧	صفر	١	صفر	٢.٢٥	صفر
-٢	-٤.٥	٩٠.٠٢	٠.٠١	٠.٩٩	٠.٠١	٢.٢٥	٠.٠٢
-١	-٣	٢٠.٠٩	٠.٠٥	٠.٩٥	٠.٠٥	٢.٢٥	٠.١١
صفر	-١.٥	٤.٤٨	٠.١٨	٠.٨٢	٠.١٥	٢.٢٥	٠.٣٤
١	صفر	١	٠.٥٠	٠.٥٠	٠.٢٥	٢.٢٥	٠.٥٦
٢	١.٥	٠.٢٢	٠.٨٢	٠.١٨	٠.١٥	٢.٢٥	٠.٣٤
٣	٣	٠.٠٥	٠.٩٥	٠.٠٥	٠.٠٥	٢.٢٥	٠.١١

من الملاحظ في الجدول أعلاه أن اقتران معلومات الفقرة يزداد كلما زاد مستوى القدرة، وأن هذه الزيادة تأخذ شكلا انسيابيا حيث تصل في أعلاها إلى القيمة (٠.٥٦) وذلك عند مستوى القدرة (ق= ١) وبعد ذلك المستوى تبدأ بالتناقص حتى تصل القيمة (١١ ٠) وذلك عند مستوى القدرة (ق = ٣)، ومن الملاحظ أن اقتران معلومات الفقرة متناسق حول النقطة (القيمة) التي تساوي مستوى صعوبتها وهي النقطة (١)، إذ نلاحظ أيضا إن اقتران المعلومات يبقى في نسق واحد وذلك في ظل النموذجين أحادي وثنائي المعلمات، إضافة إلى أن هناك علاقة بين معلمة التميز

ومقدار المعلومات حيث أنه إذا كانت معلمة التمييز للفقرة متوسطة أو قليلة تكون قيمة مقدار المعلومات قليلة أو متدنية. أي أن العلاقة بينهما علاقة طردية

ثالثاً: النموذج ثلاثي المعلمات.

ورد في فصل سابق أن النموذج ثلاثي المعلمات لا يتمتع بالخصائص الرياضية للاقتران اللوغريتمي، ولذلك فان المعادلة الخاصة بحساب مقدار معلومات الفقرة في ظل هذا النموذج يكتنفها شيء من التعقيد، وهي تأخذ الصيغة التي تتضمنها المعادلة (٥٨)

$$\text{م}_{\text{ف}(\text{ق})} = \text{ت}^{2} \left[\frac{\text{ح}_{\text{ع ف}(\text{ق})}}{\text{ح}_{\text{ص ف}(\text{ق})}} \right] \left[\frac{(\text{ح}_{\text{ص ف}(\text{ق})} - \text{تخ})^{2}}{1 - \text{تخ}} \right] \quad \ldots\ldots (58)$$

حيث،

م ف(ق) : مقدار معلومات الفقرة

ت : معلمة التمييز.

تخ : معلمة التخمين.

ح ص ف(ق): احتمال الإجابة على الفقرة إجابة صحيحة.

ح ع ف(ق) : احتمال الإجابة على الفقرة إجابة خاطئة.

ومن المهم إن ننتبه إلى إن احتمال الإجابة على الفقرة إجابة صحيحة تحسب من المعادلة (٢-٣)

$$\text{ح}_{(\text{ق})} = \text{تخ} + (1 - \text{تخ}) \left(\frac{1}{1 + \text{ث}^{-\text{ت}(\text{ق} - \text{ص})}} \right) \quad \ldots\ldots (2-3)$$

دار جرير للنشر والتوزيع

ولتوضيح كيفية استخدام المعادلة (٦ ـ ٥) سيتم حل مثال بحيث يتم حساب مقدار المعلومات لفقرة ضمن قيم معلماتها المختلفة .

مثال (٣)

فقرة معامل صعوبتها (١) ومعامل التميز لها (١.٥) ومعامل التخمين (٠.٢) فما هو مقدار المعلومات التي تقدمها عن المفحوصين في كل مستويات القدرة (− ٣ ـ ٣)؟

الحل :

سنجد مقدار المعلومات عند مستوى القدرة (−٣) وباقي الإجراءات تطبق على بقية المستويات حسب الخطوات التالية

أولاً: سنجد احتمال الإجابة على الفقرة أي سنجد أطراف المعادلة (٦ ـ ٥)

ح ف(ق) = (تخ + ١ − تخ) (١ / (١ + ث − ت(ق − ص)))

ح ف(ق) = (٠.٢) + (١ − ٠.٢) (١ / (١ + ٢.٧١٨ − ١.٥(−٣ − ١)))

ح ف(ق) = (٠.٢) + (٠.٨) (١ / (١ + ٢.٧١٨ ٦))

ح ف(ق) = (٠.٢) + (٠.٨) (١ / (١ + ٤٠٣.١٧))

ح ف(ق) = ٠.٢ + ٠.٨ (١ / ٤٠٤.١٧)

ح ف(ق) = (٠.٢) + (٠.٨) (٠.٠٠٢٤٧)

ح ف(ق) = (٠.٢) + (٠.٠١٩٧)

ح ف(ق) = ٠.٢٠١٩

- ح غ(ق) = ١- ح س(ق)

ح غ(ق) = ١ - ٠.٢٠١٩

ح غ(ق) = ٠.٧٩٨

- $\frac{١}{٠.٢٠١٩}$ = ٣.٩٥٠

- ت ٢ = (١.٥) ٢ = ٢.٢٥
- (ح س(ق) – تغ) ٢

(٠.٢٠ - (٠.٢٠) ٢ = صفر

- (١ – تغ) ٢

(١ – ٠.٢)٢

(٠.٨)٢ = ٠.٦٤

والآن سنجد مقدار المعلومات عند مستوى القدرة (ق = -٣)

$$م ف(ق) = (١.٥)^{٢} \left[\frac{٠.٧٩٨}{٠.٢٠١٩}\right] \left[\frac{صفر - ٠.٢^{٢}}{١ - ٠.٢}\right]$$

$$م ف(ق) = (١.٥) ٢ \left[\frac{٠.٧٩٨}{٠.٢٠١٩}\right] * صفر$$

= صفر

وأما بقية القيم اللازمة لحساب مقدار معلومات الفقرة في مستويات القدرة الأخرى فهي كما في الجدول (١٤).

جدول (١٤) : قيم المعلمات والقيم اللازمة لحساب اقتران معلومات الفقرة في المثال (٣)

ق	- ل	١ + ث ل	ح ص	خ م	ح ص / خ م	ح ص - تخ	مع
-٣	-٦	٣.٤٨	٠.٢٠	٠.٨٠	٣.٩٥	صفر	صفر
-٢	-٤.٥	٩٠.٩٧	٠.٢١	٠.٧٩	٣.٧٨٥	صفر	٠.٠٠٠١
-١	-٣	٢١.٠٨	٠.٢٤	٠.٧٦	٣.٢٠٢	٠.٠٠١	٠.٠١٦
صفر	-١.٥	٥.٤٨	٠.٣٥	٠.٦٥	١.٩٨	٠.٠٢١	٠.١٤٢
١	صفر	٢	٠.٦٠	٠.٤٠	٠.٦٦٧	٠.١٦٠	٠.٣٧٥
٢	١.٥	٠.٢٢٣	٠.٨٢	٠.١٥	٠.١٧١	٠.٤٢٨	٠.٢٥٧
٣	٣	٠.٠٤٩	٠.٩٦	٠.٠٤	٠.٠٤٠	٠.٤٨١	٠.٠٨٢

إن شكل اقتران المعلومات في هذا النموذج يشبه إلى حد كبير شكل الاقتران في النموذج ثنائي المعلمات عندما تكون (ص= ١، ت = ١.٥) .وعلى العموم فان المستوى العام لقيم مقادير المعلومات تعتبر متدنية، فعلى سبيل المثال عندما كان مستوى القدرة (ق = صفر) كان مقدار المعلومات للفقرة (١٤٢ .٠) وذلك في ظل النموذج ثلاثي المعلمات و(٣٤ .٠) في ظل النموذج ثنائي المعلمات، وذلك عند نفس قيم معلمات الصعوبة والتمييز، إضافة إلى أنه وفي كلا النموذجين فان الحد الأعلى لاقتران المعلومات لم يقع على مستوى القدرة الذي يلتقي (يساوي) معلمة الصعوبة، إنما وقع على مستوى أعلى من مستوى الصعوبة، والسبب في ذلك أن وجود الحدود (١- تخ) و (ح ص(ق) – تخ) في المعادلة (٦ ـ ٥) تجعل من مقدار المعلومات في ظل النموذج ثلاثي المعلمة أقل منه في نموذج ثنائي المعلمات على الرغم من تساوي قيم كل من الصعوبة والتمييز

وإذا كانت قيمة (تخ = صفر) فان مقدار المعلومات ولنفس المعلمات في ظل النموذجين سيكون متساوي، أما إذا كانت (تخ > صفر) فان مقدار المعلومات في ظل

دار جرير للنشر والتوزيع

النموذج ثلاثي المعلمات سيكون اقل منه بالنسبة للنموذج ثنائي المعلمات ولذلك فان اقتران المعلومات لفقرة في ظل النموذج ثنائي المعلمات يحدد الحد الأعلى لمقدار المعلومات في ظل النموذج ثلاثي المعلمات، وهذا شيء منطقي لان إجابة الفقرة إجابة صحيحة بالتخمين لا يضيف شيئا إلى مستوى القدرة الذي يتم تقديره

حساب اقتران معلومات الاختبار.

تم تعريف مقدار معلومات الاختبار في المعادلة رقم (٦ ــ ٢) على أنه مجموع مقادير معلومات الفقرات عند مستويات القدرة المختلفة، وبما إن إجراءات حساب معلومات الفقرة قد تم توضيحه في ظل النماذج الثلاثة لمنحنى خصائص الفقرة، فان اقتران معلومات الاختبار يمكن إيجاده أيضا، ولتوضيح ذلك سنقوم بحل مثال توضيحي في ظل النموذج ثنائي المعلمات.

مثال (٤)

تم تطبيق اختبار مكون من خمس فقرات تقيس التحصيل الأكاديمي في وحدة الكسور فاذا كانت قيم معلمات هذه الفقرات كما هي في الجدول أدناه بين كيف يمكن حساب مقدار معلومات الاختبار عند كل مستوى من مستويات القدرة.

الفقرة	الصعوبة	التمييز
١	١-	٢
٢	٠.٥-	١.٥
٣	صفر	١.٥
٤	٠.٥	١.٥
٥	١	٢

الحل:

سيتم حساب مقدار المعلومات للفقرات جميعها عند مستوى واحد هو (ق = - ٣) بالتفصيل أما للفقرات الأخرى وعند جميع المستويات فيرد دون التفاصيل في الجدول (٦ ــ٤) وعلى النحو التالي:

معطيات الفقرة (ق = −٣) (ص= −١)، (ت = ٢)

سنجد احتمال الإجابة الصحيحة على الفقرة حسب المعادلة (٢ ــ١) عند جميع المستويات .

$$\text{ح ص ف (ق)} = \frac{١}{١ + \text{ث} - \text{ت (ق} - \text{ص)}}$$

$$\text{ح ص ١ (ق)} = \frac{١}{١ + ٢.٧١٨ - ٢ (-٣ - -١)}$$

$$\text{ح ص ١ (ق)} = \frac{١}{١ + ٢.٧١٨ - ٤ ١}$$

$$\text{ح ص ١ (ق)} = \frac{١}{١ + \text{ث} - \text{ت (ق} - \text{ص)}}$$

$$\text{ح ص ١ (ق)} = \frac{١}{١ + ٥٧ ٥٤}$$

$$\text{ح ص ١ (ق)} = \frac{١}{٥٧ ٥٥} = ٠١٧ ٠ ٠$$

والآن نحسب احتمال الإجابة الخاطئة من المعادلة (ح ع ل(ق) = ١ − ح ص ١ (ق)

١ ـ ٠,٠١٧٩٩٥٣ = ٠,٩٨٢٠٠٤٧

م ل١ (ق) = ت ^٢^ ح ص (ق) * ح ع (ق)

= (٢)^٢^ (٠,٠١٧٩٩٥٣) (٠,٩٨٢٠٠٤٧)

= (٠٧١ ٠ ٠)

والآن سنكرر العملية للفقرة الثانية

ح ص ٢ (ق) = ١ / (١ + ث – ت (ق – ص))

ح ص ٢ (ق) = ١ / (١ + ٢.٧١٨ – ١.٥ (–٣ – –٠.٥))

ح ص ٢ (ق) = ١ / (١ + ٢.٧١٨ ٣.٧٥)

ح ص ٢ (ق) = ١ / (١ + ٤٢ ٥٠) =

ح ص ٢ (ق) = ١ / ٤٣.٥٠ = ٠.٠٢٢٩٨

ح ص ٢ (ق) = ١ / ٥٥.٥٧ = ٠ ٠١٧

والآن نحسب احتمال الإجابة الخاطئة من المعادلة (ح خ ٢ (ق) = ١ – ح ص ٢ (ق)

١ – ٠.٠٢٢٩٨ = ٠.٩٧٧٠

م ٢ (ق) = ث ^{٢} ح ص (ق) ● ح خ (ق)

= (١ ٥) ^{٢} (٠.٠٢٢٩٨) (٠.٩٧٧٠)

= (٠.٠٥١) .

والآن سنكرر العملية للفقرة الثالثة

ح ص ٣ (ق) = ١ / (١ + ث – ت (ق – ص))

ح ص ٣ (ق) = ١ / (١ + ٢ ٧١٨ – ١ ٥ (– ٣ – صفر)

$$\text{ح ص ٣ (ق)} = \frac{١}{١ + ٢.٧١٨^{\,٤.٥}}$$

$$\text{ح ص ٣ (ق)} = \frac{١}{١ + ٨٩.٩٧} =$$

$$\text{ح ص ٣ (ق)} = \frac{١}{٩٠.٩٧} = ٠.٠١٠٩٩٢٦$$

والآن نحسب احتمال الإجابة الخاطئة من المعادلة (ح ع د(ق) = ١ − ح ص د(ق)

١ − ٠.٠١٠٩٩٢٦ = ٠.٩٨٩٠٠٧٣

$$\text{م د(ق)} = \text{ث}^{٢}\ \text{ح ص(ق)} * \text{ح ع(ق)}$$

$$= (١.٥)^{٢}(٠.٠١٠٩٩٢٦) * (٠.٩٨٩٠٠٧٣)$$

$$= (٠.٠٢٤).$$

والآن سنكرر العملية للفقرة الرابعة

$$\text{ح ص ٤ (ق)} = \frac{١}{١ + \text{ث} - \text{ت (ق − ص)}}$$

$$\text{ح ص ٤ (ق)} = \frac{١}{١ + ٢.٧١٨ - ١.٥\ (-٣ - ٠.٥)}$$

$$\text{ح ص ٤ (ق)} = \frac{١}{١ + ٢.٧١٨^{\,٥.٢٥}}$$

$$\text{ح ص ٤ (ق)} = \frac{١}{١ + ١٩٠.٤٦} =$$

$$\text{ح ص ٤ (ق)} = \frac{١}{١٩١.٤٦} = ٠.٠٠٥٢٢٣٠٢$$

والآن نحسب احتمال الإجابة الخاطئة من المعادلة (ح ع د (ق) = ١ – ح س ٤ (ق)
١ - ٠.٠٠٥٢٢٣٠٢ = ٠.٩٩٤٧٧٦٩

م د ١ (ق) = ث ٢ ح س (ق) * ح ع (ق)
= (١.٥)٢ (٠.٠٠٥٢٢٣٠٢) * (٠.٩٩٤٧٧٦٩)
= (٠.٠١٢) .

والآن سنكرر العملية للفقرة الخامسة

$$ح\ ص\ ٥\ (ق) = \frac{١}{١ + ث - ت\ (ق - ص)}$$

$$ح\ ص\ ٥\ (ق) = \frac{١}{١ + ٢.٧١٨ - ٢(- ٣ - ١٠)}$$

$$ح\ ص\ ٥\ (ق) = \frac{١}{١ + ٢.٧١٨\ ٨}$$

$$ح\ ص\ ٥\ (ق) = \frac{١}{١ + ٢٩٧٨.٤٨} =$$

$$ح\ ص\ ٥\ (ق) = \frac{١}{٢٩٧٩.٤٨} = ٠.٠٠٠٣٣٥$$

والآن نحسب احتمال الإجابة الخاطئة من المعادلة (ح ع د (ق) = ١ – ح س ٥ (ق)
١ - ٠.٠٠٠٣٣٥ = ٠.٩٩٩٦٦٤٣

م د ١ (ق) = ث ٢ ح س (ق) * ح ع (ق)
= (٢)٢ (٠.٠٠٠٣٣٥) * (٠.٩٩٩٦٦٤٣)
= (٠.٠٠١)

والآن سنجد مقدار المعلومات للاختبار المكون من ال(٥) فقرات السابقة وذلك بجمع معلومات الفقرات الخمس وذلك حسب المعادلة رقم (٦ ــ ٢)

م (ق) = مجموع م(١ ـــ ٥)(ق)

= م١ + م٢ + م٣ + م٤ + م٥

= ٠.٠٠١ + ٠.٠١٢ + ٠.٢٤ + ٠.٠٥١ + ٠.٠٧١

= ٠.١٥٩

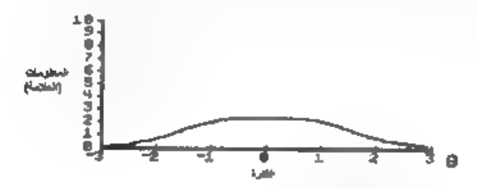
وهكذا يتم حساب احتمالات الإجابة على الفقرات الخمس عند كل المستويات ومن ثم مقدار المعلومات لكل منها وجمعها لإيجاد معلومات الاختبار عند كل مستوى من مستويات القدرة، وقد يجد البعض صعوبة في الحسابات، حيث إن الحسابات جميعها تجري بشكل حاسوبي، والآن سنورد الحلول النهائية غير المفصلة للفقرات والمستويات الأخرى في الجدول (١٥) حيث سترد مقادير المعلومات لكل الفقرات عند جميع مستويات القدرة وكذلك مقادير معلومات الاختبار عند كل مستوى من مستويات القدرة .

جدول (١٥) قيم المعلمات والقيم اللازمة لحساب اقتران معلومات الاختبار المثال (٤)[*]

مستوى القدرة	ف١	ف٢	ف٣	ف٤	ف٥	معلومات الاختبار م خ (ق)
-٣	٠.٠٧١	٠.٠٥١	٠.٠٢٤	٠.٠١٢	٠.٠٠١	٠.١٥٩
-٢	٠.٤٢	٠.١٩٤	٠.١٠٢	٠.٠٥١	٠.٠١٠	٠.٧٧٧
-١	١	٠.٤٩٠	٠.٣٣٦	٠.١٩٤	٠.٠٧١	٢.٠٩١
صفر	٠.٤٢	٠.٤٩٠	٠.٥٦٣	٠.٤٩٠	٠.٤٢٠	٢.٣٨٣
١	٠.٠٧١	٠.١٩٤	٠.٣٣٦	٠.٤٩٠	١	٢.٠٩١
٢	٠.٠١٠	٠.٠٥١	٠.١٠٢	٠.١٩٤	٠.٤٢٠	٠.٧٧٧
٣	٠.٠٠١	٠.٠١٢	٠.٠٢٤	٠.٠٥١	٠.٠٧١	٠.١٥٩

[*] تم حساب البيانات في خلايا الجدول بنفس الطريقة التي تمت على الفقرة الأولى لجميع مستويات القدرة وهي عادة ما يتم حسابها من خلال برامج احصائية خاصة مثل برامج (ROM)

من الملاحظ أن اقترانات (مقادير) معلمات الفقرة تشكل نسقا قبل وبعد (حول) مستوى القدرة الذي يساوي قيمة معلمة صعوبة تلك الفقرة، وكما يلاحظ ان قيمة التمييز بشكل عام تدور حول القيمة(١.٥) وبالنسبة للصعوبة فكانت تدور حول القيمة (صفر) وتبعا لذلك فان معلومات الاختبار أيضا أخذت نسقا عددا فوق وتحت مستوى القدرة التي تساوي مستوى صعوبتها، وتبعا لذلك نتج شكل اقتران المعلومات للاختبار كما يظهر في الشكل (٦ ـ ٤) حيث يظهر أن مقدار المعلومات مسطح نسبيا وذلك ضمن مدى القدرة (-١ ـ١) أما خارج هذا المدى فان مقدار المعلومات يزيد ويقل بشكل متناوب وبنمط متسق، وهذا يشير الى أن مقدار المعلومات التي يوفرها لاختبار تختلف من مستوى قدرة الى مستوى قدرة آخر انظر الشكل (٣٤)



شكل(٣٤) شكل منحنى اقتران المعلومات لاختبار مكون من خمس فقرات الواردة في المثال رقم 4

تفسير اقتران معلومات الاختبار.

إن ما يثير الانتباه في شكل اقتران معلومات الاختبار هو اعتماده على الهدف أو الغرض الذي صمم من اجله، فمن الممكن إطلاق التفسيرات العامة في هذا الصدد ومنها: أن اقتران معلومات الاختبار يبلغ أعلى قيمة له عند عدة نقاط متصل السمة، حيث لا تكون الدقة في هذه الحالة مختلفة من نقطة (مستوى) إلى أخرى، وهذا يكون أفضل وضع للاختبار من حيث تقديره لقدرة المفحوصين الذين يقعون على أو حول مستوى القدرة الذي يقابل أعلى قمة لشكل اقتران المعلومات .

دار جرير للنشر والتوزيع

في بعض الاختبارات يكون شكل اقتران المعلومات مسطحا على معظم نقاط (مستويات) القدرة على متصل القدرة وتبعا لذلك يكون مستوى الدقة، وفي بعض الاختبارات تكون المعلومات دقيقة لكنها ضمن مدى معين من القدرة، ومن الضروري أن يكون الاختبار مميز (جذاب)* للمفحوصين في المستويات التي تقابل قمة شكل الاقتران وما حولها ولذلك فانه يجب الانتباه عند تفسير اقتران المعلومات إلى العلاقة التبادلية بين كل من مقدار المعلومات ومدى التغير في تقديرات القدرة، وذلك لتوضيح مقدار المعلومات وتفسيره من خلال الخطأ المعياري للتقدير، ومن جل ذلك لا بد من الانتباه إلى إن الجذر التربيعي لأعلى مقدار من المعلومات يعكس الحصول عليه، حيث نلاحظ أعلى مقدار من مقادير المعلومات عند مستويات القدرة المختلفة، ويمكن تقدير الخطأ المعياري للتقدير من المعادلة (٥٩)

$$خ_{م ت} = \sqrt{\frac{1}{\text{أعلى مقدار للمعلومات}}} \quad \ldots\ldots\ldots (٥٩)$$

وعلى سبيل المثال في الشكل (٦ ـ ٤) فان أعلى قيمة لاقتران لمعلومات الاختبار هي (٢.٣٨٣) وهي عند مستوى القدرة (ق = صفر) لاحظ جدول (٦ ـ٤) وهذا يتم ترجمته الى خطأ معياري مقداره حسب المعادلة (٦ ـ ٦)

$$خ_{(م ت)} = \sqrt{1 / \text{أعلى مقدار للمعلومات}}$$

$$= 1 / ٢.٣٨٣$$

$$= ٠.٦٥$$

وهذا يعني أن ما نسبته (٦٨٪) من التقديرات عند هذا المستوى من تلك القدرة تقع بين (- ٠.٦٥ ـ ٠.٦٥) وهذا يعني أن تقديرات هذا المستوى تمت بشكل مقبول من حيث الدقة

* جذاب هنا تعني القدرة على التمييز وهنا يتطلب الاهتمام بالبدائل او المموهات

خلاصة

- إن المستوى العام لاقتران معلومات الاختبار يعتمد على ثلاثة عوامل :
  - عدد الفقرات المكونة للاختبار.
  - معدل معلمات التمييز لفقرات الاختبار
  - النموذج اللوغريثمي المستخدم لمنحنى خصائص الاختبار
- شكل منحنى اقتران معلومات الاختبار يعتمد على عاملين.
  - توزيع معلمات الصعوبة على متصل القدرة.
  - توزيع المعدل العام لمعلمات التمييز للفقرات
- عندما تتجمع صعوبات الفقرات حول قيمة معينة فينتج عن ذلك قمة لمنحنى الاقتران على متصل القدرة، ويعتمد الحد الأعلى لهذه القمة على قيم معلمات التميز.
- عندما تتوزع صعوبات الفقرات بشكل (مدى) واسع فان اقتران معلومات الاختبار سيميل إلى ان يكون مسطحا أكثر مما لو كانت تتوزع ضمن مدى ضيق أو محدد
- إذا كانت قيم معلمة التمييز (< ١) فان قيمة أو مقدار اقتران معلومات الاختبار تميل إلى أن تكون قليلة
- إذا كانت قيمة معلمة التمييز ت(> ١.٧) فان قيمة المعلومات الناتجة تميل إلى أن تكون أعلى ما يمكن.
- في ظل النموذج ثلاثي المعلمات فان قيم معلمات التخمين (تخ) تكون أكثر من (صفر) وهذا يقلل من مقدار المعلومات للاختبار وذلك في مستويات القدرة المتدنية، إضافة إلى ان ارتفاع قيمة معلمة التخمين بشكل عام يقلل من مقدار

معلومات الاختبار، وكأننا نطرح قيمة معلومة مقدرة من معلومات الاختبار ولكن ليس نتيجة للمعرفة الحقيقية إنما نتيجة للتخمين.

- من الصعوبة الحصول على منحنى اقتران أفقي (خط أفقي) لمعلومات الاختبار، وللحصول على ذلك لا بد من انتشار معلمات الصعوبة على مدى واسع على متصل القدرة، إضافة إلى ضرورة إن تكون قيم معلمات التمييز تتوسط الحد الأدنى لمدى التمييز وتتخذ توزيع على شكل حرف (U).

الفصل السابع

تدريج الاختبار وتحديد خصائصه

Test Calibration

إن الحديث عن مفهوم تدريج الاختبار يتطلب الوعي بكل من مفهوم منحنى خصائص الفقرة ومنحنى خصائص الاختبار واقتران المعلومات للفقرة والاختبار، وكذلك مواصفات متصل القدرة، فمن المفترض مثلا أن يؤخذ بعين الاعتبار أن متصل القدرة تتوسطه القيمة(صفر) كما إن وحدة القياس فيه هي الواحد الصحيح، ويمتد ضمن المدى النظري (- ∞ ـــ ∞)، كما أن قيم تقديرات معلمات الفقرات والمفحوصين تقع على المتصل ضمن المدى المشار إليه، حيث تعتبر هذه الافتراضات كمقدمة نظرية لمفاهيم النظرية الحديثة في القياس، لكن هذه المقدمة لا تكفي لوحدها لتوضيح وبيان واقع العملية الاختبارية، فباني الاختبار عليه أن يكون على وعي بما تقيسه الفقرات التي يقوم بصياغتها وما هو مستوى القدرة الذي تتناسب معه هذه الفقرات (متدني متوسط، عالي) إذ أن التحديد المسبق لقيم معلمات الفقرات أمر غير ممكن، إضافة إلى أن التحديد المطلق (الحاسم) للمستوى الحقيقي لقدرة المفحوصين وبمجرد تطبيق الاختبار على مجموعة من المفحوصين أمر من الصعب تحقيقه.

ومن هنا فان من المهام الرئيسية لباني الاختبار هي تحديد قيم معلمات الفقرات وقدرات المفحوصين على متصل القدرة، ويشار إلى هذه الإجراءات عادة بعملية تدريج الاختبار (Test Calibration) حيث تقدم هذه العملية إطارا مرجعيا لتفسير نتائج الاختبار. وتتم عملية التدريج هذه من خلال تطبيق الاختبار على مجموعة من المفحوصين بحيث تأخذ إجاباتهم على فقرات الاختبار نمطا ثنائيا(صح خطأ أو ١، صفر) وتجري مجموعة من العمليات الرياضية للبيانات المستمدة من إجابات المفحوصين على فقرات الاختبار، وذلك من خلال البرامج الإحصائية

بالحاسب الآلي، للحصول على متصل القدرة (الفريد) وبشكل عملي، وموحد لكل من الفقرات والمفحوصين، بعد ذلك يتم تقدير معلمات الفقرات والمفحوصين على متصل القدرة حيث أن تدريج الاختبار(قدرة، صعوبة، تمييز) يعتبر جزء مهم من الجهازا لاختبار ووضعه قيد الاستخدام، وهذا يدعم تفسير وتوضيح البناء النظري لنظرية القياس الحديثة.

إجراءات تدريج الاختبار

كان آلن بيرنبوم (Alen,Bernbaum,1968) هو أول من اقترح فكرة تدريج الاختبار في عام (١٩٦٨)، بعد ذلك انتشر هذا المفهوم بشكل واسع واستخدمت فيه البرامج الحاسوبية مثل برمجية بيسكال (Biscal) لرايت وميد (Wright and Mad,1976) وبرمجية لوجست (Logist) لونجرسكي وبارتون (Wingersky and Barton and gord,1982) وتستخدم عملية التدريج هذه مفهوم أعلى نسبة ترجيح "Maximum Liklihood" .

ففي المرحلة الأولى من عملية التدريج يتم تقدير معلمات فقرات الاختبار وفي المرحلة الثانية يتم تقدير قدرة المفحوصين، حيث تتفاعل (تتداخل) كل من هاتين المرحلتين مع بعضهما البعض، وذلك من اجل الحصول على التقدير النهائي للمعلمات، وبعد هذه المرحلة فان الاختبار يكون قد تم تدريجه، وكذلك يتم تعريف متصل القدرة (السمة) أو تحديده من حيث قيم القدرة الواقعة عليه ومواقع المفحوصين عليه وصعوبة وتمييز الفقرات.

وحسب إجراءات بيرنبوم (Bernbaum) يتم تقدير القدرة للمفحوصين في أكثر من خطوة حتى نصل إلى أقرب تحديد أو تقدير للسمة المقاسة (القدرة)، وبعد ذلك يتم تقدير معلمات كل فقرة في الاختبار، وذلك حسب الإجراءات التي أشرنا إليها في الفصول السابقة، حيث تجري هذه الإجراءات لكل الفقرات، وذلك بسبب استقلالية الفقرات عن بعضها البعض

بعد ذلك تبدأ المرحلة الثانية وعلى افتراض أن قيم التقديرات التي تم الحصول عليها في المرحلة الأولى هي أفضل تقدير لمعلمات الفقرات، وتتم هذه الخطوة بتقدير

قدرات المفحوصين كما ورد في سابقا حيث يفترض أن قدرة المفحوصين مستقلة عن بعضها البعض، لذلك يتم تقدير قدرة كل مفحوص على حدة، ويمكن القول إن الإجراءات التي أشرنا إليها في الفصل الثالث من الباب الثاني من هذا الكتاب تمثل المرحلة الأولى من عملية تدريج الاختبار، بينما إجراءات الفصل الخامس تمثل المرحلة الثانية من عملية التدريج. إذ يتم تكرار الإجراءات في المرحلتين حتى نحصل على أفضل تقدير للقيم التي تحقق المحك المطلوب، ولا بد هنا من الإشارة إلى تأثير التداخل في المرحلتين حيث تتم إجراءات التدريج لكل من الفقرات والمفحوصين في وقت واحد، وللتقليل من تعقيدات العمليات الحسابية وتعددها فلا بد من إجراءها حاسوبيا.

مشكلة وحدة القياس

تعتبر وحدة القياس من أهم المشكلات التي واجهت اقتراح بيرنبوم اذ أن هذا الاقتراح لم يقدم وحدة قياس متفردة (وحيدة) لمتصل القدرة، حيث أن نقطة المتصف ووحدة القياس على متصل القدرة يكتنفها بعض الغموض، فالقيم المختلفة قد تبدو متساوية وبمعنى آخر فان متصل السمة لا بد أن يقبل التحويل الخطي، وكنتيجة لذلك فانه من الضروري أن تصاغ بعض القواعد المشتركة في تحديد نقطة المتصف لوحدة القياس لمتصل القدرة، وهذا يطرح تساؤلا من قبل المختص بالحاسوب لتنفيذ إجراءات بيرنبوم آليا.

فمثلا في برمجية بيسكال (BICAL) فان هناك عملية تتم بعد مرحلة التدريج الأولى تسمى الفقرات المشتركة (Anchor) ولذلك فان كلا المرحلتين وضمن عملية الفلترة تتم بوحدة قياس مختلفة، وفي الوقت الذي تقترب فيه عمليات الفلترة من بعضها البعض فان ذلك يؤدي إلى اقتراب وحدة القياس لمتصل السمة واقتراب نقطة المتصف العملية، وكذلك وحدة القياس. إذ أن المهمة الرئيسية في هذه الإجراءات هي الوصول إلى وحدة قياس تعتمد على عينة الفقرات المكونة للاختبار وكذلك استجابات عينة المفحوصين على الاختبار ككل لأنه من غير الممكن الحصول على تقديرات حقيقية لقدرات المفحوصين ومعلمات الفقرات للسمة المقاسة، وأن أفضل

إجراء هو الحصول إلى وحدة قياس تعتمد على مركب من المفحوصين والفقرات المكونة للاختبار (قاسما مشتركا لكل من المفحوصين والفقرات) وهذا ما يسمى بعملية (Anchoring) وهي مجموعة من الفقرات المشتركة بين مختلف مستويات القدرة وتسمى أحيانا بفقرات المرساة .

تدريج الاختبار في ظل النموذج احادي المعلمات (نموذج راش).

عرف في النظرية الحديثة للقياس ثلاثة نماذج لمنحنى خصائص الفقرة ومن الممكن استخدام أي منها، إضافة إلى عدة طرق لتطبيق نموذج (عرض) بيرنبوم، ومن بين هذه النماذج تم اختيار مدخلا يعتمد على النموذج الأدائي الذي يعرف عادة باسم صاحبه راش (Rasch Model) حيث قام بنجامين رايت (Benjamin Wright) بتطبيقها وذلك من خلال برمجية بيكال (BICAL) وفي ظل هذا النموذج يتم تقدير معلمة واحدة لكل فقرة، وتتم إجراءات التدريج في ظل هذا النموذج لعدد قليل نوعا ما لكل من الفقرات والمفحوصين ومن هنا فان عدد الفقرات المشتركة أو فقرات المرساة (Anchor) يكون قليل وتبعا لذلك تكون إجراءات عملية التدريج ابسط إذا ما قورنت في النماذج الأخرى .

ولتوضيح إجراءات التدريج سنورد مثالا يبين إجراءات التدريج وذلك لاختبار مكون من (١٠) فقرات تم تطبيقه على (١٦) مفحوص، حيث تمت هذه الإجراءات بشكل حاسوبي حيث كانت الفقرات تتناسب مع معدل قدرة المفحوصين ككل وقد كانت استجابات المفحوصين على الفقرات تأخذ واحد من درجتين، الدرجة (١) للإجابة الصحيحة و (صفر) للإجابة الخاطئة، وسيكون الهدف من هذا المثال هو استخدام استجابات المفحوصين لتدريج الاختبار، حيث سنلاحظ القيم المناظرة (Vector) لكل فقرة وكيفية اشتقاق هذه القيم من خلال إجابة جميع المفحوصين على الفقرات كل على حده واستجابة المفحوصين ككل وتظهر تلك القيم في الجدول (١٦)

جدول ١٦: استجابات (١٦) مفحوص على اختبار مكون من (١٠) فقرات وكذلك القيم الناظرة.

	رقم المفحوص	الفقرات										الدرجة الكلية
		١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	
	١	١	٠	١	٠	٠	٠	٠	١	٠	٠	٢
	٢	١	٠	١	٠	٠	٠	٠	٠	٠	٠	٢
	٣	١	١	١	٠	١	٠	١	٠	٠	٠	٥
	٤	١	١	١	٠	١	٠	٠	٠	٠	٠	٤
E	٥	٠	٠	٠	٠	١	٠	٠	٠	٠	٠	١
X	٦	١	١	٠	١	٠	٠	٠	٠	٠	٠	٣
A	٧	١	٠	٠	٠	٠	١	١	١	٠	٠	٤
M	٨	١	٠	٠	٠	١	١	٠	٠	١	٠	٤
I	٩	١	٠	١	٠	٠	١	٠	٠	١	٠	٤
N	١٠	١	٠	٠	٠	١	٠	٠	٠	٠	١	٣
E	١١	١	١	٠	١	١	١	١	١	١	١	٩
E	١٢	١	١	١	١	١	١	١	١	١	٠	٩
	١٣	١	١	١	٠	١	٠	١	٠	٠	١	٦
	١٤	١	١	١	١	١	١	١	١	١	٠	٩
	١٥	١	١	٠	١	١	١	١	١	١	١	٩
	١٦	١	١	١	١	١	١	١	١	١	١	١٠

عرفنا في الفصل الخامس أنه لا يمكن تقدير قدرة أي مفحوص إذا أجاب على جميع فقرات الاختبار إجابة صحيحة أو لم يجب على جميع فقرات الاختبار إجابة صحيحة، بمعنى أنه إذا حصل على علامة كاملة أو علامة (صفر)، وإذا نظرنا إلى الجدول (١٣) فان المفحوص رقم (١٦) سيتم استبعاده من إجراءات التدريج، كما عرفنا أن الفقرة التي يجيب عليها جميع المفحوصين أو تلك التي لم يجب عليها أي من

المفحوصين سيتم استبعادها من الاختبار وبالتالي فان كل مفحوص أو فقرة تنطبق عليهما الشروط السابقة لن تدخل البيانات الخاصة بكل منهما في إجراءات عملية التدريج، وإذا ما عدنا إلى الجدول السابق سنجد أن هذا الحديث ينطبق على المفحوص رقم (١٦) بالنسبة للمفحوصين الذين طبق عليهم الاختبار، ولم ينطبق على أي فقرة بالنسبة للفقرات المكونة للاختبار.ومن إجراءات نموذج راش (Rasch) أيضا أن المفحوصين الذين يجيبون على نفس العدد من الفقرات إجابة صحيحة سيحصلون على نفس المستوى من القدرة، حيث أنه ليس من الضروري التمييز بين هؤلاء المفحوصين واللذين سيحصلون على نفس العلامة الخام، وهكذا سيتم استخدام الفقرات التي أجيب عنها بشكل فردي وكل ما نحتاجه هو عدد المفحوصين ضمن أي علامة خام من العلامات التي تم الحصول عليها نتيجة لتطبيق الاختبار، أما المفحوص الذي أجاب على الفقرة إجابة خاطئة فسوف لن يهتم به وبسبب عدم الاهتمام هذا تم حذف الفقرة التي اجيب عليها إجابة خاطئة من جميع المفحوصين، وتنفيذا لإجراءات تدريج الاختبار فان البيانات التي سنحصل عليها نتيجة للإجراءات التي وردت سابقا ستكون على الشكل الذي ستظهر عليه في جدول (١٧)

جدول (١٧): علامات المفحوصين والفقرات الواردة في جدول (١٦) حسب إجراءات نموذج راش.

العلامة	الفقرات										العلامة الكلية
	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	
١	٠	٠	٠	٠	١	٠	٠	٠	٠	٠	١
٢	١	٠	٢	٠	٠	٠	٠	١	٠	٠	٤
٣	٢	١	٠	١	١	٠	٠	٠	١	٠	٦
٤	٤	١	٢	٠	٢	٣	١	١	٢	٠	١٦
٥	١	١	١	٠	١	٠	١	٠	٠	٠	٥
٦	١	١	١	٠	١	٠	١	٠	٠	١	٦
٩	٤	٤	٢	٤	٤	٤	٤	٤	٤	٢	٣٦
المجموع	١٣	٨	٨	٥	١٠	٧	٧	٦	٧	٣	٧٤

حيث تمثل الصفوف العلامات الخام التي جاءت ضمن المدى (١ ـــ ٩) أما الصف الهامشي الأسفل فيمثل مجموعة الإجابات الصحيحة التي أبداها المفحوصين على الفقرات، أي علامات الفقرات إذا جازت التسمية، أما العمود الأخير في الهامش الأيسر فيمثل مجموع الإجابات (العلامات) لجميع المفحوصين في كل علامة خام تم الحصول عليها، وللتوضيح فان مجموع العلامات في العلامة الخام (١) هو (١) وحصل من إجابة الطالب رقم (٥) على الفقرة رقم (٥) وكذلك مجموع العلامات في العلامة الخام (٤) هو (١٦) وحصل من إجابات المفحوصين ذوي الأرقام (٤، ٧، ٨، ٩) على فقر ت مختلفة وكذلك مجموع إجابات المفحوصين الذين حصلو على العلامة (٩) هو العلامة (٣٦) وهكذا .

واستكمالا للحل فان البيانات التي سوف يتم استخدامها في ظل نموذج راش وطريقة بيرنبوم هي فقط العمود الأيسر والصف السفلي في الهامش، وهي من ميزات نموذج راش حيث الاختصار للبيانات والعمليات الحسابية باستخدام أعلى نسبة ترجيح، خاصة في ظل تداخل الإجراءات كونها تعتمد على بيانات كل من المفحوصين والفقرات في الاختبار. ومن أجل استكمال إجراءات التدريج يتم أخذ عوامل (Vectore) التكرارات ومنها يتم الحصول على تقديرات صعوبة الفقرات بما في ذلك الفقرات المشتركة (Anchor Items) وتوضيعها على متصل السمة كمرحلة أولى، وفي ظل نموذج راش ومن خلال الفقرات المشتركة يتم توظيف مبدأ عام هو أن معلمة التميز للفقرة يثبت عند القيمة (١) لجميع الفقرات، ولذلك فان وحدة القياس لتقدير قدرات المفحوصين مثبت عند القيمة (١) وكل ما تبقى هو تحديد منتصف متصل السمة، وفي إجراءات برنامج (BICAL) الكمبيوتري يتم تحديد نقطة المنتصف لمتصل السمة من خلال متوسط تقديرات صعوبات الفقرات، وللحصول على قيمة مقبولة لنقطة المنتصف يتم اشتقاقها من قيم صعوبة الفقرات ككل، وذلك من خلال تكرار هذه العملية حتى الحصول على قيمة صعوبة مساوية للقيمة (صفر) . وبسبب توحيد كل من وحدة الصعوبة والقدرة فان نقطة المنتصف ووحدة القياس النهائية لا بد من تحديدهما أيضا. ويتم ذلك بين مرحلتي التدريج، حيث أن تقديرات القدرة التي

تمت في المرحلة الأولى ستستخدم في تحديد نقطة منتصف متصل السمة من خلال إعادة تقدير معلمات الفقرات التي تم لحصول عليها في مرحلة التدريج الأولي، حيث سيتم تقدير القدرة لكل علامة خام في الاختبار، والتي حصلنا عليها في الإجراء الثاني الذي تتلخص نتائجه في الجدول (١٤) حيث تستخدم صعوبات الفقرات بعد الفلترة كما لو كانت معلمات الصعوبة والتي تأتي من خلال مجموع علامات الفقرات (مجموع الاستجابات على الفقرات).

إن نتائج هذه المرحلة ستمثل تقديرات القدرة لكل علامة خام وردت في الجدول (١٣) وهنا يتم التأكد من صحة الإجراءات من خلال برمجية (BICAL) وقد لخص رايت (Wright) القيمة المطلقة للفروق بين قيم تقديرات صعوبة الفقرات لمرتي الفلترة من خلال معيار محدد مفاده أنه إذا كان مجموع تقديرات الصعوبة للفلترة أقل من (٠٠١) فإن عملية التقدير تكون قد انتهت، أما إذا كان مجموع القيم المطلقة لتقديرات الصعوبة للفلترة أكبر من القيمة (٠٠١) فلا بد من متابعة الفلترة، بمعنى أن مرحلتي التدريج سيتم إعادتهما، ولذلك فإن إجراءات المرحلة الأولى التي تتضمن عملية (Anchor) ستستمر في المرحلة الثانية للتأكد من الحصول على قيمة مقبولة للوصول إلى المحك المطلوب وهو القيمة الأقل من (٠٠١)، وعند تحقق ذلك فإن القيمة الناتجة لتقديرات معلمات القدرة ستكون هي المعتمدة، وسيكون متصل السمة قد تم تحديده فعلا، ومن خلال الإجراءات الكمبيوترية فإن تقديرات معلمات الصعوبة للفقرات في المثال الوارد في الجدول (١٣) ستؤول إلى الصورة المبينة في الجدول (١٥) .

جدول ١٨: تقديرات معلمة الصعوبة للفقرات الواردة في جدول (١٦)

الفقرة	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
الصعوبة	٢.٣٧-	٠.٣٧-	٠.٣٧-	٠.٩٨	١.٠٠--	٠.١١	٠.١١	٠.٥٢	٠.١١	٢.٠٦

يبين الجدول قيم التقديرات لمعلمة صعوبة فقرات الاختبار، ويمكننا التأكد من أن مجموع معلمات الصعوبة للفقرات يساوي = (صفر) مع الأخذ بعين الاعتبار أخطاء التقدير، وهكذا فان مجموع الصعوبات السالبة هو: (-٢.٣٧+-٠.٢٧ + + -٠.٢٧ + + - ١= (٣.٩١) ومجموع الصعوبات الموجبة هو: (٠.٩٨ + ٠.١١ + ٠.١١ + ٠.١١ + ٠.١١ + ٥٢٠ + ٢.٠٦ = ٣.٨٩) وهما قيمتين متساويتين تقريبا خاصة إذا ما أخذنا بعين الاعتبار أن بعض القيم مقربة، وقد مر معنا في الفصل الثاني تقدير معلمات الفقرة حيث أن صعوبة الفقرة رقم (١) هو القيمة (- ٣٧ ٢) وقد وضعت هذه القيمة الفقرة على الحد الأدنى من متصل السمة وكذلك الفقرة رقم (٦) التي بلغت قيمة الصعوبة لها (٠.١١) ووضعتها على منتصف متصل السمة تقريبا، أما الفقرة رقم (١٠) فقد وصلت قيمة معلمة الصعوبة لها القيمة (٢.٠٦) وهذه القيمة تضعها على الحد الأعلى من متصل السمة ، وعادة ما يتم تفسير معلمة الصعوبة للفقرة من خلال موقعها على متصل السمة، حيث تمثل القيم التي يتم الحصول عليها متوسط صعوبة الاختبار بالنسبة للمفحوصين وذلك بسبب الإجراءات التي تتضمنها عملية الـ (Anchoring) حيث أن فقرات المرساة (المشتركة) تكون اقل من مستوى بعض المفحوصين وأعلى من البعض الآخر.

وعلى الرغم من الحصول على تقديرات القدرة لكل مفحوص كما ستتبين في الجدول (١٦) إلا أن كل المفحوصين اللذين حصلوا على نفس العلامة الخام سيحصلون على نفس التقدير، فالمفحوصين رقم (١، ٢) حصلوا على علامة خام بلغت (٢) ولذلك حصلوا على نفس التقدير وهو القيمة(١.٥) وكذلك المفحوصين (٧، ٨، ٩) حصلوا على العلامة الخام (٤) ولذلك حصلوا على نفس التقدير وهو القيمة (-٤٢ ٠) .

جدول ١٩: العلامات الخام والتقديرات المقابلة لها من خلال بيانات الجدول (١٦)

رقم المفحوص	العلامة الخام	تقدير القدرة	رقم المفحوص	العلامة الخام	تقدير القدرة
١	٢	١.٥ -	٩	٤	٠ ٤٢ -
٢	٢	١.٥ -	١٠	٣	٠ ٩١ -
٣	٥	٠ ٠٢	١١	٩	٢ ٣٣
٤	٤	٠.٤٢ -	١٢	٩	٢ ٣٣
٥	١	٢ ٣٧ -	١٣	٦	٠ ٤٦
٦	٣	٠ ٩١ -	١٤	٩	٢ ٣٣
٧	٤	٠.٤٢ -	١٥	٩	٢ ٢٣
٨	٤	٠.٤٢ -	١٦	١٠	———

وهذا ما يعكس حقيقة ثبات قيمة معلمة التمييز عند القيمة (١) لكل الفقرات المكونة للاختبار في ظل نموذج راش، وهذا يعزز الادعاءات بأن نموذج راش يعتريه بعض الغموض ويحاجة إلى الدفاع عنه في بعض الأحيان، إذ أن ثبات معامل التمييز ادعاء يعتريه بعض الغموض، وذلك عندما يحصل جميع المفحوصين ممن حصلوا على نفس العلامة الخام على نفس التقدير، أي أن علامتهم تساوي عدد الفقرات التي أجابوا عليها إجابة صحيحة، بينما يختلف الوضع في حال استخدمنا النموذجين ثنائي وثلاثي المعلمات، حيث لا تعتمد علامة المفحوصين على عدد الفقرات التي أجابوها فقط بل تعتمد كذلك على نمط الاستجابات التي يبديها المفحوصين على فقرات الاختبار، وفي ظل هذين النموذجين فان المفحوصين الذين يجيبون على نفس العدد من الفقرات ونفس نمط الاستجابة سيحصلون على نفس التقدير لمستوى قدراتهم، بمعنى أن النموذجين الثنائي والثلاثي المعلمات يضيفا شرطا جديدا لتساوي تقديرات مجموعة من المفحوصين جابوا نفس العدد من الفقرات ألا وهو أن يكون هؤلاء المفحوصين أجابوا نفس الفقرات وذلك بسبب اختلاف صعوبة الفقرات

ومن هنا فانه إذا تساوى عدد من المفحوصين في عدد الفقرات التي أجابوها وكانت هذه الفقرات مختلفة فإنهم لن يحصلوا على نفس تقدير القدرة بسبب إجابتهم على فقرات مختلفة الصعوبة رغم تساوي عددها، وبمعنى أكثر تحديدا فان العلامات الخام التي حصل عليها المفحوصين بسبب تساوي عدد الفقرات التي أجابوا عليها في حقيقة الأمر تعبر عن قدرات أجابت على فقرات مختلفة الصعوبة، لان الفقرات متوسطة الصعوبة مثلا تتطلب مستوى قدرة متوسط والفقرات الصعبة والفقرات الأكثر صعوبة تتطلب مستوى قدرة أعلى وبملاحظة الجدول (١٣) نجد أن عدد المفحوصين الذي دخل في الحسابات (١٥) بسبب استبعاد المفحوص رقم (١٦) بسبب إجابته على جميع الفقرات وهذا من متطلبات إجراءات التدريج حسب نموذج راش كما بينا سابقا، إذ أن التقديرات تضع المفحوص على موقع يطابق قدرته فالمفحوص رقم (٧) مثلا حصل على تقدير علامة خام (٤) وتقدير (- ٠.٤٢) وهذا وضعه على موقع (نقطة) أقل من نقطة منتصف متصل القدرة وهكذا بالنسبة لبقية المفحوصين، كما يمكن رسم شكل التوزيع لتقديرات القدرة والحصول على الإحصائيات الخاصة بها، ففي المثال الحالي بلغ المتوسط الحسابي لتقديرات القدرة (٠.٠٦) كما بلغ الانحراف المعياري لها(١ ٥٧) ولذلك فالمفحوص رقم (٧) مثلا حصل على تقدير ينحرف بمقدار (٢٧ ٠) انحراف معياري تحت المتوسط الحسابي لان الانحراف سالب، وعلى كل حال فان احد لن يستوعب تفسير علامات القدرة للمفحوصين من خلال توزيع علامات مجموعة المفحوصين، ولاستيعاب وفهم ذلك فانه لا بد من الوعي بأن تفسير علامات القدرة لا يكون من خلال العلامات مباشرة وكما يعكسها موقع المفحوص على متصل القدرة، حيث أن الحصول على التقديرات مر بإجراءات حسابية معقدة وفيها تكرار للحصول على وضع ما ملائم لافتراضات النظرية والنموذج اللوغريتمي المستخدم وكل ذلك يتم حاسوبيا

ملخص إجراءات تدريج الاختبار

إن النتيجة النهائية لعمليات تدريج الاختبار تتمثل في تحديد متصل القدرة وفي ظل نموذج راش (Rasch) فان وحدة القياس لمتصل السمة هي القيمة (١) وله نقطة

متصف تتمثل في النقطة (صفر) كما أن الأفكار التي وردت في الفصل السادس تكاد تكون سطحية وغير متعمقة، ولكنها لا تعكس بالتفصيل كيفية تحديد متصل السمة، حيث يعتمد هذا التفصيل على استجابات المفحوصين على الفقرات المكونة للاختبار والتي تعكس حالة مركبة من خصائص المفحوصين وخصائص الفقرات وهذا ما قام بيرنبوم (Bernbum) باختراقه من خلال نموذجه الذي اعتمد على نموذج راش. حيث أنه وبما أن القيم الحقيقية لمقياس السمة غير محددة فان المقياس الذي افترضه بيرنبوم (Bernbum) يستخدم كما لو كان هو المقياس الحقيقي، وفي ظل هذا المقياس يتم تفسير كل من قيم الصعوبة والقدرة للمفحوصين وكذلك تدريج الاختبار معا، حيث ينتج عن عملية التدريج تحديد مواقع المفحوصين والفقرات على مقياس القدرة، ففي المثال السابق فان الفقرة رقم (٥) مثلا حصلت على معلمة صعوبة بلغت قيمتها (- ١) والمفحوص رقم (١٠) حصل على تقدير قدرة بلغت قيمته (- ٠.٩١) ولذلك فان احتمال إجابة هذا المفحوص على الفقرة رقم (٥) هو (٠.٥) وبذلك فان إمكانية تحديد الفقرات والمفحوصين على مقياس مشترك يعتبر من نقاط القوة لنظرية السمات الكامنة، حيث يسمح ذلك بتفسير نتائج عملية تدريج الاختبار ضمن إطار واحد وتوفير معنى واحد لتقديرات المعلمات.

خلاصة

- في تدريج الاختبار فان البيانات التي نحصل عليها نتيجة لعمليات التدريج ستختلف باختلاف عدد المفحوصين الذين يجيبون على جميع الفقرات أو لا يجيبوا على أي من فقرات الاختبار وكذلك بالنسبة للفقرات التي أجيب عليها من قبل جميع المفحوصين (السهلة) أو تلك التي لم يجب عليها أي مفحوص (الصعبة) باعتبار أن بياناتها لا تدخل في إجراءات التحليل وذلك لأن هذا الحذف سيؤثر في البيانات التي نحصل عليها والتي تؤثر بالتالي في عمليات التدريج.
- بما أن الاختبارات يتم تصميمها بحيث تعد سهلة أو متوسطة الصعوبة أو صعبة فان نتائج عملية التدريج لا بد أن تعكس ذلك بالضرورة، ويعود ذلك إلى وجود الفقرات المشتركة (Anchor) في الاختبار حيث أن وجود ثلاث حالات للاختبار سيؤدي إلى أن يكون متوسط صعوبة الاختبار مساويا للصفر وهي قيمة منتصف متصل السمة
- في كل عملية تدريج فان المفحوصين الذين يحصلون على نفس العلامة الخام، سوف يحصلون على نفس تقدير القدرة في ظل نموذج راش لكن ذلك لا يحدث في ظل النموذجين ثنائي وثلاثي المعلمات.
- إذا خضعت مجموعة من المفحوصين لعدة اختبارات فان المتوسطات الحسابية والانحرافات المعيارية لتقديرات قدراتهم ستختلف في كل عملية معايرة لكل اختبار وهناك عدة عوامل تساهم في هذا الاختلاف منها :
  - اجراءات الفقرات المشتركة (Anchoring) والتي تقرب متوسط تقديرات القدرة من متوسط صعوبة فقرات الاختبار حيث يؤدي ذلك إلى أن يكون متوسط صعوبة الاختبار السهل قيمة موجبة، ومتوسط صعوبة الاختبار الصعب قيمة سالبة ومتوسط صعوبة الفقرات المشتركة يقترب من الصفر

- عدد المفحوصين حيث يمكن أن يقل عدد المفحوصين بسبب إجراءات التدريج التي تتطلبها عملية التدريج حيث الحذف للمفحوصين أو للفقرات ويزداد اثر ذلك إذا كان عدد المجموعة قليل في الأصل، وتبعا لذلك فان مقدار المعلومات في ظل النماذج اللوغريتمية الثلاثة يميل إلى أن يكون قليلا، ولذلك ليس من الضروري أن تمثل تقديرات قدرة المفحوصين والتي يتم الحصول عليها مستويات قدراتهم الحقيقية، لأنها في افضل الأحوال هي تقدير لتلك القدرة .

- إن إجراءات الفقرات المشتركة (Anchoring) تدفع بقيمة متوسط صعوبة الفقرات إلى الاقتراب من الصفر وبالتالي إلى تحديد قيمة منتصف متصل السمة إلى القيمة (صفر)، إذ أن النتيجة المباشرة لذلك هي العلامة لمتوسطة (صفر) على الاختبار في حالاته الثلاث (صعب سهل، متوسط) وفي كل مرة من مرات التدريج رغمن منحنى خصائص الفقرة ويعود ذلك بالطبع إلى أن معلمة الصعوبة لجميع الفقرات تساوي (١)

- على الرغم من أن بعض مقادير اقتران المعلومات للاختبار في حالاته الثلاث قد تكون متشابهة كما في الشكل (٧ ــ ٢) لكن ثمة فروقا جوهرية بينهما، فالمنحنى الذي يمثل الاختبار المشترك (الفقرات المشتركة) يتشابه مع منحنى الاختبار السهل رغم انه مسطحا أكثر منه، وهذا مؤشر على أن الاختبار حافظ على مستوى الدقة خلال مدى واسع من القدرة، أما منحنى اقتران المعلومات للاختبار الصعب فقد حقق مقدارا قليلا من المعلومات حيث وصلت قيمته إلى نقطة المنتصف أي انخفض مقدار ما قدمه عن المفحوصين من ذوي القدرات العالية أو يلاحظ انحداره الملحوظ عند مستوى القدرة (ق = صفر) إذا ما قورن بمنحنى اقتران المعلومات للاختبارين(١، ٢) إضافة إلى أنه لم يحقق مستوى الدقة المرغوب به للمعلومات.



تحديد خصائص الاختبار

Specifying the Characteristics of A Test

هناك العديد من الاختبارات التي يتم تصميمها وفقا للنظرية الكلاسيكية في القياس وقد تم تحليلها اعتمادا على النظرية الحديثة في القياس، وهذا يشكل حالة من عدم الانسجام بين منهجية البناء وآلية التحليل مما يقلل من إبراز نقاط القوة للنظرية الحديثة (IRT) ولذلك فانه ومن أجل فهم النظرية الحديثة بميزاتها لا بد من بناء وتصميم وتحليل وتفسير الاختبارات اعتمادا على الإطار النظري لهذه النظرية، ومن هنا فان الغرض الرئيسي هنا هو تزويد القارئ بمفاهيم واضحة تتضمن تقنيات واضحة حول بناء الاختبار اعتمادا على الإطار النظري لنظرية القياس الحديثة

إن عملية بناء وتصميم الاختبارات تطورت إلى الدرجة التي أصبحت فيها عمل مؤسسي يقوم به أشخاص مختصون لأنها - الاختبارات - أصبحت تستخدم على نطاق واسع، حيث تستخدم من قبل الشركات والمصانع والمؤسسات الحكومية والمؤسسات التربوية خاصة في ظل ضبط جودة التعليم (Quality of Educational Process) إضافة إلى استخدامها من قبل المعلمين لقياس التحصيل الأكاديمي، وتعتمد كل هذه الهيئات في عملها على إجراءات علمية من اجل انتقاء مجموعة من الفقرات من مجتمع كبير من الفقرات لتشكيل الاختبار، إما عملية الاختبار هذه فتتم اعتمادا على مجموعة من المعايير الخاصة بالفقرات كالمحتوى والخصائص السيكومترية أو بناء على المعلمات الخاصة بها، وحسب النظرية الحديثة يتم تجميع الفقرات ذات الخصائص المرغوبة حسب المعايير المعتمدة ويحتفظ بها لحين الاستخدام، وذلك ضمن ما أصبح يعرف ببنك الأسئلة (Item Bank) كي نستطيع الحصول على أي عينة من الفقرات (اختبار) التي تمتلك خصائص يحددها مستخدم الاختبار حسب الصعوبة والتمييز والقدرات المرغوبة، وإذا تم ذلك فانه يمكن اختيار الفقرات من المجتمع الكبير

من الفقرات (بنك الأسئلة) حيث يمكن معرفة خصائص الفقرات (صعوبة تمييز) قبل تطبيقها على مجموعة من المفحوصين على عكس ما هو مألوف في النظرية الكلاسيكية حيث لا يمكن معرفة خصائص الفقرات إلا بعد تطبيقها على عينة المفحوصين وهي مختلفة حسب مجموعة التطبيق، وفي النظرية الحديثة فانه لا بد أن تتطابق خصائص الفقرات مع الأهداف المحددة للاختبار، وإذ لم تتطابق فانه لا بد من استبدال الفقرات حتى نصل إلى الخصائص المرغوبة التي تحقق تطابق الفقرات مع أهداف الاختبار، وبذلك فانه يمكن تجنب الكلفة المكررة نتيجة لبناء اختبارات في كل مرة استخدام، وعندما نتحدث عن الكلفة فإننا نعني المبالغ المالية التي يتم إنفاقها والجهد المبذول والوقت المستغرق في البناء والتطبيق والتصحيح والتحليل والتفسير، وللحصول على عينة من الفقرات (اختبار) فانه من الضروري مراعاة الخطوات التالية :

- تحديد السمة الكامنة التي سيقيسها الاختبار .
- صياغة الفقرات لقياس هذه السمة.
- تطبيق الفقرات لاستبعاد الفقرات الضعيفة منها
- اختيار مجموعة الفقرات التي تقيس السمة قيد الاهتمام
- تطبيق الفقرات على مجموعة كبيرة من المفحوصين
- اختيار نموذج منحنى خصائص الفقرة المراد تصميم الاختبار من خلاله.
- تحليل استجابات المفحوصين على الفقرات لتدريج الاختبار

بعد ذلك تعتبر نتائج عملية التدريج هي المقياس الأساسي لفقرات الاختبار، ومن خلال الأدب الخاص ببناء الاختبار فانه يمكننا معرفة أو تحديد فقرات معروفة من حيث قيم معلماتها وهذا ما يسمى بالتدريج المسبق للفقرات

تطوير اختبار بالتدريج المسبق

بما أن الفقرات قبل عملية التدريج تصمم لقياس سمة محددة فان اختيار اختبار من هذه الفقرات فلا بد أن يقيس أيضا نفس السمة، وقد يبدو أن ذلك بحاجة إلى القليل من الوقت، حيث ان هناك عدة أسباب تفرض الحاجة إلى اختبارات إضافية

لقياس نفس السمة، ومنها على سبيل المثال وجود نماذج بديلة من الاختبار من اجل المحافظة على سريته، إضافة إلى أن وجود عدة طبعات من الاختبار يمكننا من استخدامها للأغراض المدرسية كالمسابقات والأنشطة والترشيحات والاختيار وفي هذه الحالات فان الفقرات يتم اختيارها من نفس المجموعة الكلية (بنك الأسئلة) لنفس السمة والتي حصلنا عليها بنفس الإجراءات المتعلقة بخصائص الفقرات التي تحقق الغرض من الاختبار

إن لعملية التدريج المسبق للفقرات عدة ايجابيات منها انه يمكن استخدام قيم معلمات الفقرات في إيجاد منحى خصائص الاختبار واقتران المعلومات وذلك قبل تطبيقه على المفحوصين، ويمكن ذلك ولا غرابة فيه فالمنحنيات تعتمد على توزيع علامات القدرة للمفحوصين على متصل القدرة، ومن هنا فان كل من منحنى الخصائص والمعلومات للاختبار يمكن الحصول عليها طالما أن معلمات الفقرات قد تم توفيرها، وهذا من شأنه إعطاء باني الاختبار فكرة مسبقة عن كيفية أداء الاختبار (كيف سيسير) قبل تطبيقه على المفحوصين، إضافة إلى أنه عندما يتم تطبيق الاختبار وتدريجه فان إجراءات المعايرة يمكن أن تتم للحصول على تقديرات القدرة لمجموعة أخرى من المفحوصين من خلال نفس المقياس لمجموعة الفقرات الأم

الأهداف العملية للاختبار

يمكن تزويد القارئ بمجموعة من أنماط الأهداف العملية للاختبار حيث يسهل ذلك على القارئ وباني الاختبار من حيث تحديد الاختبار المراد تصميمه، إضافة إلى انه يسهل من الوعي بالإجراءات الحاسوبية التي تتم لتحليل وتفسير نتائج الاختبار، وسوف ترد هذه الأنماط من خلال التعرف على أنواع الاختبارات من وجهة نظر النظرية الحديثة في القياس، وهي على النحو التالي:

أنواع الاختبارات

أولاً: اختبارات التصفية (الغربلة) Screening tests.

وهي اختبارات تهدف إلى التمييز بين المفحوصين ممن هم فوق او تحت مستوى قدرة محدد بهدف تصفيتهم لاختيار الأفضل من بينهم، وتستخدم مثل هذه الاختبارات عادة لتحديد المنح والبعثات الدراسية للطلبة أو للتوجيه نحو برامج تعليمية محددة لأغراض علاجية أو التعيين في المراكز المهمة التي تحتاج لقدرات عالية

ومن المعلومات التي من المهم الوعي بها بالنسبة لهذه الاختبارات ما يلي :

- إن المنحنى المرغوب لمنحنى خصائص اختبار التصفية هو منحنى تتطابق فيه نقطة منتصف مدى القدرة الحقيقية مع علامة القطع المحددة، حيث يكون المنحنى عميق ما أمكن عند ذلك المستوى من القدرة.
- يشكل اقتران معلومات الاختبار فجوة في أقصى حد له عند مستوى القدرة الذي يساوي قيمة نقطة القطع.
- تتجمع قيم معلمات صعوبة الفقرة حول علامة القطع المحددة، وفي أحسن الأحوال تكون قيم صعوبات الفقرات متطابقة مع قيم علامة القطع، أما قيم معلمات التمييز للفقرة فتكون عالية، وهذا بالطبع غير واقعي بسبب تساوي عدد من الفقرات من حيث قيم الصعوبة لها، حيث يمكننا اختيار مجموعة من الفقرات التي توفر أعلى مقدار من المعلومات عند نقطة القطع التي تتساوى مع قيمة معلمة الصعوبة لها، حيث أن الفقرة تقدم أعلى مقدار من المعلومات عند مستوى القدرة الذي يتساوى مع قيمة معلمة الصعوبة لها.

ثانيا : الاختبارات واسعة النطاق Broad - ranged tests.

وهي اختبارات تستخدم لقياس قدرة المفحوصين فوق حد معين من القدرة حيث يكون غالبيتهم ضمن مدى محدد وتستخدم لأغراض المقارنة بين المفحوصين ممن يقعون فوق المعيار المحدد، وذلك لإصدار أحكام او أوصاف حولهم من اجل إجراء



المقارنات بينهم مثل اختبارات القراءة والرياضيات ذات المحتوى الواسع .ومن المعلومات التي من المهم الوعي بها بالنسبة لهذه الاختبارات ما يلي :

- إن المنحنى المرغوب لمنحنى خصائص اختبار واسع النطاق هو منحنى تتطابق فيه نقطة منتصف مدى القدرة الحقيقية مع علامة منتصف مدى القدرة قيد الاهتمام والذي يكون فيه مستوى القدرة مساويا للصفر(ق = صفر) حيث لا بد أن يكون منحنى خصائص الاختبار خطيا في معظم المدى الذي يمثله.
- إن اقتر ن لمعلومات المرغوب به للاختبار يكون خطيا كلما كان مدى مستويات القدرة واسعا، وأما أعلى قيمة للمعلومات فتكون عند الحدود العليا للنطاق.
- تنتشر قيم معلمات صعوبة الفقرة بشكل منتظم فوق مستوى القدرة العملي، حيث ينزع اقتر ن المعلومات إلى الخط الأفقي كلما ركزنا على الحصول على معلومات أكبر، وللحصول على اقتران معلومات افقي فلا بد أن تتوزع صعوبات الفقرات ذات التميز المتدني على شكل حرف (U) ، وكلما أبدت الفقرات مقدار معلومات متدني أثر ذلك على دقة المعلومات حيث تنزع إلى الانخفاض .

ثالثا: اختبارات القمم (الفجوات) Peaked tests .

وهي اختبارات تشكل حالة الوسط بين اختبارات التصفية واختبارات واسعة النطاق بمعنى أنها تستخدم لقياس القدرة البسيطة على نطاق واسع لكن هذا النطاق أضيق من نطاق اختبار واسع النطاق، وتستخدم كذلك في الحالات التي يكون فيها عدد المفحوصين كبيرا وينقسمون إلى مجموعات متمايزة أي يوجد فجوات في توزيع قدراتهم، وبمعنى أكثر تحديدا أنهم ينقسمون إلى مجموعات متمايزة من حيث القدرة. ومن المعلومات التي من المهم الوعي بها بالنسبة لهذه الاختبارات ما يلي

- إن المنحنى المرغوب لخصائص اختبار القمم هو منحنى تتساوى فيه نقطة منتصف مدى القدرة الحقيقية مع منتصف مدى القدرة قيد الاهتمام، حيث يجب ان يكون ميل المنحنى متوسط عند ذلك المستوى من القدرة.

- ان اقتران المعلومات المرغوب يجب أن تصل أعلى قيمة له عند نقطة منتصف مدى العلامات الحقيقية على منحنى خصائص الاختبار نفسه، كما يبدأ اقتران معلومات الاختبار بالدوران عند الحدود العليا للقدرة قيد الاهتمام
- تبدو قيم معلمات صعوبة الفقرة على شكل عناقيد تتجمع حول نقطة منتصف مدى القدرة، لكنه ليس بنفس الكيفية التي يبدو فيها بالنسبة لاختبارات التصفية، ولا بد لقيم معلمات التمييز أن تكون عالية وبشكل واضح، كما لا بد أن تكون قيم معلمة التمييز للفقرات التي تقع صعوبتها ضمن مدى القدرة أعلى منها للفقرات التي تقع صعوبتها خارج ذلك المدى

وأخيرا فان النظرية الحديثة تطرح إطارا نظريا آخر لقياس قدرة المفحوص وعلاقة معلمات المفحوصين ومعلمات الفقرات المكونة للاختبار، وان فهم هذا الإطار بافتراضاته المتعددة يتطلب من المهتمين وخاصة من طلبة ودارسين القياس والتقويم الاطلاع والدراسة حول هذه النظرية أكثر وأكثر وما هذا الكتاب إلا إطارا أوليا ولكنه أساسي لفهم النظرية الحديثة في القياس.

الباب الرابع

مقارنة بين النظرية التقليدية والنظرية الحديثة

الفصل الأول

مقارنة بين النظرية الكلاسيكية والنظرية الحديثة للقياس

النظريات والنماذج

عادة ما تستخدم نظرية الاختبار و نموذج الاختبار بشكل مرادف على الرغم من وجود فروق بينهما فنظرية الاختبار تقدم إطارا عاما يربط المتغيرات الملاحظة **كدرجة الاختبار ودرجة الفقرة**، مع متغيرات غير ملاحظة مثل **الدرجة الحقيقية ودرجة القدرة**. وبسبب عمومية النظرية لا يمكن القول بأنها مفيدة بشكل تام أو عديمة الفائدة، لذا فان هذه النظرية تتضمن عدة مفاهيم مثل الدرجة الحقيقية ودرجة الاختبار والدرجة الخطأ ولذلك لا يمكن الحكم عليها بالفائدة أو عدمها إلا بعد وضعها في نماذج خاصة توضح إجراءاتها، وقد تم وضع هذه النماذج في إطار نظرية الاختبار ووصفت في اختبارات مفصله للعلاقات ضمن مجموعة من مفاهيم نظرية الاختبار والتي تعتمد على مجموعة الافتراضات المتعلقة بمفاهيم الاختبار وعلاقتها بهذه النماذج علما بان ملائمة هذه النماذج يعتمد على طبيعة البيانات.إضافة إلى أن هذه التقسيمات تتوفر من خلال الأبحاث التجريبية أو نماذج من دراسات متطابقة والتحليل المنطقي لافترضات النموذج بالنسبة لبيانات الاختبار تؤدي إلى صدق النموذج

فالاختبار الذي يحتوي فقرات من نوع الاختيار من متعدد وبسبب التخمين الناتج عن الصعوبة يقوم (الاختبار) على افتراض بان الدرجة الحقيقية والدرجة الخطأ غير مرتبطتان وهذا ربما لا يطبق وربما يكون مسؤولا عن انخفاض الدرجة الخطأ للمفحوصين من ذوي القدرات العليا وارتفاع الدرجة الخطأ للمفحوصين من ذوي القدرات الدنيا ومثل هذه النتيجة ستشكل انتهاكا للافتراضات

إن بعض النماذج عادة ما تقدم تمثيلا غير متكاملا لمجموعة البيانات المتطابقة لذلك فان الحال سيكون على نحو غير متطابق لذلك فان السؤال الملح هو كون النموذج صحيح أم غير صحيح؟ عندما نشعر أنها تقدم تمثيلا غير متكامل للبيانات فالنماذج الكلاسيكية من الاختبارات توصف بأنها نماذج ضعيفة لان الافتراضات لهذه النماذج تحول بسهولة باستخدام بيانات الاختبار مثلا النماذج ذات الحد بين التي تعتمد على افتراضات تقتصر على توزيع العلامات الخطأ تعتبر نماذج قوية ونماذج الاستجابة على الفقرة تعتبر نماذج قوية كذلك لان افتراضاتها متشددة وتجعل إمكانية تحويل البيانات وتغيرها أمرا صعبا فالنماذج ذات المعلم أو المعلمين أو الثلاثة معالم تقوم على افتراضات قوية وتقول بان مجموعة الفقرات التي يتكون منها الاختبار تقيس خاصية أو سمة واحدة فقط بينما النماذج التقليدية لا تعتمد على مثل هذه الافتراضات ومن المهم أن نفترض أن قواعد البناء (Factor Structure)أيا كانت على أنها نماذج متوازية .

اهمية نظريات القياس ونماذجها

إن لكل من هذه النظريات والنماذج الخاصة بها أهمية خاصة في التطبيقات التربوية وفي القياس النفسي لأنها تقدم إطارا لاعتماد النتائج ومواجهة المشاكل واهم هذه النتائج هي معالجة أخطاء القياس، وتعتبر النظرية أو النموذج جيدين بالقدر الذي تساهما فيه في فهم آلية التأثير التي تنتج عن هذه الأخطاء وذلك من خلال:

- حساب قدرة الفرد وكيفية تخفيض مساهمة الخطأ مثل lengthening test .
- الارتباطات بين المتغيرات مثل (disattenuation formulas) .
- تقرير الدرجة الحقيقية أو درجة القدرة وفترة الثقة الخاصة بكل منهما

لقد عالجت هذه النظريات الأخطاء بطرق مختلفة فاعتبار الأخطاء على أنها توزيعا طبيعيا على نموذج واحد مثلا لا يعني توزيعا افتراضيا للأخطاء على نموذج آخر ففي النموذج الواحد فان حجم أخطاء القياس يمكن اعتبارها ثابتة على مقياس

علامات الاختبار (الخطأ المعياري للقياس) وفي مكان آخر يمكن اعتبار حجم الأخطاء على أنها ترتبط بالعلامة الحقيقية كما في نموذج الأخطاء ذو الحدين وتعين الأخطاء في النموذج له تأثير نوعي على كيفية حساب العلامة الخطأ وتقديرها .

إن النظرية أو النموذج، الجيد يمكن أن يقدم مرجعية لتصميم الاختبار أو حلول لبعض المشاكل العملية، كذلك يحدد دقة العلاقة بين فقرات الاختبار وعلامة القدرة ، حيث يؤدي ذلك إلى الحصول على توزيع مرغوب للعلامات، فالأخطاء في العينة يمكن تحملها فالاختبارات في الحاسب فان النماذج التي تربط بين حساب القدرة وإحصائيات الفقرة تحتاج إلى آلية في انتقاء الفقرات فيمكن انتقاءها في أي نقطة ضمن إجراءات الاختبار التي توفر اكبر قدر من المعلومات عن قدرة المفحوص وهذا يعني أن المفحوصين على مقياس عام (كما في نماذج الاستجابة) مما يؤدي إلى اختيار بشكل يوفر المعلومات الأكثر فائدة حول القدرة .

فقد أشار لورد (Lord .Fredric,1952) قبل ٤٠ سنة إلى فكرة مهمة مفادها أن مفهوم كلا من الدرجة الملاحظة والدرجة الحقيقية (True Score) للمفحوص لا تترادفان مع مفهوم القدرة (ability) وذلك لاستقلال هذه الدرجة عن الاختبار، في حين أن الدرجتين الظاهرية والحقيقية غير مستقلتان عن الاختبار بمعنى أن المفحوصين يتقدموا إلى الاختبار وهم على مستويات معينة من القدرة التي يقيسها الاختبار، الأمر الذي يعني حصول المفحوصين على علامات حقيقية متدنية على الاختبارات الصعبة وعلامات حقيقية مرتفعة على الاختبارات السهلة لكن علامات القدرة تبقى ثابتة على أي اختبار يبنى لقياس محتوى معين، وبمرور الزمن فان هذه القدرة تتغير بفعل التعليمات وعوامل اخرى ولكن عند التقدير فان لكل طالب القدرة تعرف بارتباطها مع المحتوى.حيث أن علامات القدرة التي تكون مستقلة عن اختيار فقرات الاختبار تكون ذات أهمية لأنها تسمح بالمقارنة معيارية المرجع

وقد اهتم لورد (lord) ومعه مجموعة من المختصين في القياس (psychometricians) بنظريات ونماذج تصف المفحوصين وصفا مستقلا عن أي اختبار يمكن تطبيقه وقد لاحظ هؤلاء أن تطبيق القياس سوف يعزز فيما لو كانت

دار جرير للنشر والتوزيع

خصائص الاختبار وفقراته تجعل العينة مستقلة، إن الخطوة الأولى في هذا الاتجاه في النظرية التقليدية هي تفضيل معامل الارتباط بايسيريال .Biserial.Corr) coefficient.) على معامل الارتباط بوينت بايسيريال .Point .bis corr) coefficient.) في تقدير تمييز الفقرة لان الأول (bias) أكثر ثباتا من الثاني (P bis) على عينة من الطلبة .

على أي حال فان الإحصائيات المتعلقة بالفقرة كالصعوبة و التميز والإحصائيات المتعلقة بالاختبار كالصدق والثبات تعتمد جميعها على عينة المفحوصين التي تم الحصول على الإحصائيات من جراء التطبيق عليها، لكن هذه ليست من القضايا الهامة لان معظم الاختبارات تم بناءها بهذه الطريقة وبقيت كذلك حتى نهاية الستينات الأمر الذي يؤكد الاهتمام بعينة المفحوصين المناسبة لتوفير فقرات ذات إحصائيات مناسبة وكذلك للتمكن من بناء اختبارات متوازية .

لقد قدم المختصين في القياس تقديرات مختلفة لإحصائيات الفقرات حيث ،رتبطت هذه التقديرات بجوانب كثيرة كالنماذج والطرق المرتبطة بنظرية الاستجابة على الفقرة وقد كان لورد أول من أثار هذه الفكرة وذلك في أطروحته لنيل درجة الدكتوراه في سنة ١٩٥٣ تحت عنوان (thought at the time the field was known as latent trait theory)

حيث يفترض باني الاختبار في المجال النفسي والتربوي عادة أن هناك سمات أو خصائص معينة يشترك فيها جميع الأفراد ولكنهم يختلفون في مقدار امتلاكهم لها، وبالرغم من أن هذه السمات غير محسوسة ولا يمكن قياسها بطريقة مباشرة إلا أنه يمكن الاستدلال على مقدارها من السلوك الملاحظ للفرد والمتمثل في استجاباته على فقرات الاختبار، وهذا ما أوجب تسمية النظرية الحديثة بنظرية السمات الكامنة، فالسمة التي تكمن وراء استجابة الفرد على فقرات اختبار لفظي مثلا، تختلف عن السمة التي تكمن وراء استجابته على فقرات اختبار مكاني أو عددي. ولكن يمكن أن تكمن سمة واحدة وراء استجابته على فقرات اختبارين مختلفين متعلقين بنفس المحتوى.

ويتمثل الفرق بين كل من النظريتين الكلاسيكية والحديثة في الافتراضات التي تستند إليها كل من النظريتين من حيث مدى امتلاك السمة قيد القياس وكذلك المفاهيم الخاصة بالاختبار وخصائصه،وتبعا لذلك التوصيف للسمة ففي النظرية التقليدية نتحدث عن الدرجة الحقيقية، وفي النظرية الحديثة نتحدث عن القدرة، والواقع أن الدرجة الحقيقية والقدرة يرمزان للشيء ذاته ولكن بمقاييس مختلفة إلا أن الفارق بينهم هو أن الدرجة الحقيقية (True Score) تعتمد على الفقرات التي يتكون منها الاختبار، في حين أن مقياس القدرة (Ability) مستقل عن الفقرات التي يتشكل منها الاختبار أما أهم الفروق بين النظريتين فهي على النحو التالي:

مفهوم الثبات مقابل مفهوم اقتران المعلومات

توفر النظرية الحديثة مفهوما جديدا يختلف عن الثبات يدعى اقتران المعلومات والذي نحصل عليه من تطبيق فقرة أو اختبار. وهذه الاقتران يعني مقدار الثقة في أننا حصلنا على معلومات تقودنا إلى تقدير قدرة فرد معين أو مجموعة من الأفر د في مستوى معين من القدرة.أي أن مقدار المعلومات نحصل عليها عند كل مستوى قدرة على طول مقياس القدرة، أو بمعنى آخر أننا نحصل على قيم متعددة لكمية المعلومات التي حصلنا عليها من البند أو الاختبار إذا هي تختلف عن الثبات الذي هو عبارة عن مؤشر و حد فقط للاختبار لكل الأفراد الذين طبق عليهم.ويمثل اقتران المعلومات عادة بمنحنى يبين أعلى وأقل مقدار معلومات حصلنا عليها عند مستويات القدرة المختلفة.

مجموعة التطبيق مقابل مقياس القدرة.

يلعب اقتران المعلومات دورا مهما في النظرية الحديثة مثل الدور الذي يلعبه الثبات في النظرية التقليدية إلا أن الثبات يعتمد على الاختبار والمجموعة التي طبق عليها الاختبار ولهذا يكون تعميمها غير ممكنا.أما اقتران المعلومات فيعتمد فقط على مقياس القدرة وعلى منحنى استجابة الفقرة، كذلك فإن الثبات مؤشر للخطأ المعياري في القياس وهذا المؤشر ثابت على جميع مستويات الدرجة الحقيقية،بينما يعتبر اقتران المعلومات مؤشر للخطأ المعياري في التقدير وهذا الخطأ يحسب عند كل مستوى من

مستويات قدرة الطالب وتعرف بالتالي دقة تقدير قدرة الفرد على كل مستوى من المستويات.

مقبولية مؤشرات الفقرات مقابل تطابق الفقرات مع النموذج.

مؤشرات الاختبار في النظرية التقليدية تتغير بتغير المجموعة التي يطبق عليها الاختبار أما مؤشرات الاختبار في النظرية الحديثة فهي ثابتة **Invariance** بصرف النظر عن المجموعة التي طبق عليها الاختبار كما أن قدرة الشخص ثابتة مهما كانت الأسئلة التي طبقت عليه . حيث يتم الاهتمام بمعامل الصعوبة والتمييز والموهات ،اعتمادا على مجموعة التطبيق التي لا بد من الاهتمام بخصائصها من حيث تجانس أفرادها وتمثيلهم للمجتمع،وكذلك تمثيل الفقرات للمحتوى موضوع القياس،في حين انه في النظرية الحديثة يتم الاهتمام بمقدار المعلومات التي تقدمها كل فقرة والمعلومات التي يقدمها الاختبار ككل ومدى مطابقته للنموذج اللوجستي المستخدم، حيث يتم اختيار الفقرات التي تناسب قدرة المفحوص حيث يقع كل من الصعوبة والقدرة على مقياس واحد.

الفصل الثاني

النظرية الكلاسيكية

هي نظرية تتحدث عن درجة الاختبار من خلال أبعاد ثلاث هي الدرجة الحقيقية، الدرجة الظاهرية والدرجة الخطأ ولها عدة نماذج منها نموذج خطي بسيط) (X=T+E ولان هناك مجهولان لكل مفحوص فالمعادلة لا تحل قبل معرفة أحد هذين المجهولين ، وتقوم هذه النظرية على الافتراضات التالية :

- الدرجة الحقيقية (ح) والأخطاء(خ) غير مرتبطتان
- متوسط الأخطاء يساوي صفر أي أن مجموع (ن) = صفر
- الدرجات الخطأ على الاختبارات المتوازية غير مرتبطة .

وفي هذه الحالة فان الدرجة الحقيقية هي الفرق بين الدرجة الملاحظة والدرجة الخطأ وتعرف العلامة الحقيقية بأنها توقع الدرجة الظاهرية على اختبارات متوازية والاختبارات المتوازية هي الاختبارات التي تقيس نفس المحتوى والتي يحصل فيها المفحوص على نفس الدرجة الحقيقية ويكون فيها خطأ القياس متساوي

وقد أصبح بالإمكان الحصول على نتائج مهمة نتجت عن نماذج هذه النظرية مثل قانون(Sperman - Brown formula) وهو العلاقة التي تربط طول الاختبار بصدقه، والمعروفة في تطوير الاختبارات، وللحصول على نماذج أخرى اتجه الباحثون نحو إسقاط بعض الفرضيات أو التعديل على البعض الآخر أو إضافة توزيعات افتراضية للأخطاء أو للدرجة الحقيقية، فإتباع توزيع الأخطاء للتوزيع الطبيعي أو توزيع ذو الحدين أصبح مألوفا حيث أن النموذج الثاني ذو الحدين يفيد في حل مشكلات المقاييس محكية المرجع ويستخدم في حساب طول الاختبار بنتائج الثبات ومستوى الإتقان . إن تعريف توازي الاختبار يعتبر ضعيفا فبدلا من تساوي العلامة الحقيقية على الاختبار استبدل بتشكيلهما علاقة خطية وما زال بعض الباحثون يعتمدون على النماذج الكلاسيكية للاختبار لتحديد العلامة الخطأ ببيان مركبات الخطأ، مثل الناتجة عن العلامة أو الناتجة عن الفقرات وتصميم الدراسات لفحص

النتائج وأثرها على تباين علامة الاختبار وثبات الاختبار. إن معظم التركيز في نظرية الاختبار الكلاسيكية ينطوي على النماذج في مستويات علامة الاختبار والذي يغاير نظرية الاستجابة على الفقرات،أي أن النموذج يربط علامة الاختبار بالعلامة الحقيقية والتي تعتبر أكثر قوة من ربط علامة الفقرة بالعلامة الحقيقية ومن هنا فان إحصائيات الفقرة مثل الصعوبة (P) والتميير(r) وصلتهما بإحصائيات الاختبار مثل الثبات والوسط والانحراف المعياري قديمة وقد استخدمت لإخراج الاختبار بخصائص مرغوبة .

وتكفي الإشارة إلى أن خصائص الفقرة تم برهنتها وفحصها وان عينات هذه الإحصائيات مستقلة وهذه الاستقلالية تجعلها ذات فائدة معينة وهذه الفائدة تنخفض إذا كانت العينة غير مماثلة للمجتمع، وقد ظهر استخدام فقرات المرسى (anchor items) في حقل الاختبارات في إدارة الاختبارات العادية والتي يمكن استخدامها في حل المشاكل المتعلقة بالعينات والتي تجعل التحليل معقدا

إن من ايجابيات نماذج النظرية الكلاسيكية أنها تعتمد على افتراضات مرنة – إن جاز التعبير – ونسبية أي انه يمكن مقارنتها بالبيانات الحقيقية كذلك فان مؤشرات المفحوص أي درجته ومؤشرات الفقرة أي الصعوبة والتميز مستقلة عن الاختبار وعن عينة الاختبار وهذه الاستقلالية تحد من التدخل الإحصائي في التطوير العملي للاختبار.

محددات النظرية الكلاسيكية :

١. الصعوبة و التميز اللذان يشكلان حجر الأساس غير مستقلان ويشكلان أساسا في تطبيق نماذج هذه النظرية ويعتمدان على عينة المفحوصين وبلغة أخرى(لغة التميز) انه يمكن الحصول على قيم أعلى من العينات غير المتجانسة وعلى قيم منخفضة من العينات المتجانسة وبلغة الصعوبة يمكن الحصول على قيم أعلى من العينات ذات القدرة المتوسطة وقيم منخفضة من

العينات ذات القدرة منخفضة

٢. العلامات التي نحصل عليها غير مستقلة داخليا لذلك فان صعوبة الاختبار تؤثر على النتائج مباشرة وهي حسب نظرية الاستجابة على الفقرة احتمالية أن يجيب المفحوص على الفقرة إجابة صحيحة، وإذا كان من الضروري تصميم اختبار ذو معايير محددة كاختبارات المرشحين للدراسة فان الاستجابة لمادة الامتحان تجيز لواضع التصميم أن يفعل ذلك وبالتحديد فان خاصية الاستجابة للمواضيع قابلة للتطبيق في التطبيقات الحديثة كالاختبارات بالحاسب ولهذه النظرية مساوئ تقنية مثل أن نماذجها معقدة وتزداد مشاكل حساب المعالم عند تطبيق النموذج و ما زال النموذج الملائم مشكلة حيث لم يحدد بعد كيفية اختياره و لمساوئ المترتبة على تطبيقه وعلى الغالب فان الحاجة التقنية تميل لان تكون أكثر تعقيدا منها في نماذج النظرية التقليدية

أوجه القصور

سيطرت نظرية القياس التقليدية على حركة القياس لفترة ليست بالقليلة،ونتيجة لبحوث والدراسات حول الدرجة ومكوناتها ومحاولة البحث عن حلول لما عجزت النظرية عن تفسيره بدا الحديث عن منحى آخر لنظرية القياس التقليدية،حيث عانت النظرية التقليدية من بعض أوجه القصور ومنها

أولا،

إن طرق تحليل الاختبارات المبنية على النظرية التقليدية والمفاهيم السيكومترية المرتبطة بها، مثل معاملات الصعوبة والتمييز وفعالية المموهات، تختلف باختلاف خصائص العينة المستخدمة في حساب هذه المعاملات . مثلا: لو سحبنا عينة مكونة من ١.. شخص وحسبنا معامل الصعوبة لبند معين وكانت قيمته ..٧٥ فهل سنحصل على نفس هذه القيمة لو سحبنا عينة أخرى مكونة من ١ شخص آخرين؟

الجواب. لا نحصل على نفس القيمة لأن العينة مختلفة. إذا فإن مؤشر الصعوبة ليس من المؤشرات المستقرة لارتباطه ارتباطا مباشرا بطبيعة العينة كما أن معامل التمييز (وهو ارتباط البند بالدرجة الكلية) يتأثر بتباين استجابة العينة، فمدى تباين الاستجابة يعكس مدى تباين العينة، وبذلك فإن معامل التمييز يرتبط بخصائص العينة، لذلك من الصعب أن نفصل قيمة معامل التميز عن طبيعة العينة إذا فإن عملية اعتمادنا على معامل الصعوبة ومعامل التميز لا تعتبر من الجوانب الجيدة نظرا لأن هذين العاملين مرتبطان إلى حد كبير بطبيعة وخصائص العينة المستجيبة، وبذلك من الممكن أن نحصل على معاملات تمييز وصعوبة مختلفة من عينة لأخرى تبعا لاختلاف العينة

ثانيا :

إن أداء الأفراد على الاختبار يختلف باختلاف بنود الاختبار التي تم سحبها من مجتمع البنود الكبير، فهذه البنود قد تختلف في صعوبتها وهذا الاختلاف في صعوبة البنود سينعكس في نهاية المطاف على الأداء المختلف للأفراد من اختبار لآخر وقد ينتفي وجود الاختلاف في أداء الأفراد إذا تضمنت الاختبارات نفس البنود أو بنود مكافئة لها من حيث مستوى الصعوبة.

ثالثا:

إن النظرية التقليدية تفترض تساوي تباين أخطاء القياس لجميع الأفراد الذين يطبق عليهم الاختبار. ولكننا نلاحظ في بعض الأحيان أن أداء بعض الأفراد يكون في الاختبار أكثر اتساقا من غيرهم من الأفراد، وأن هذا الاتساق يختلف باختلاف مستوى قدرتهم وهذا فإننا ربما نتوقع أن الأفراد من ذوي القدرة المرتفعة يكون أداؤهم في صورة موازية من صور الاختبار أكثر اتساقا من أداء الأفراد ذوي القدرة المتوسطة . فمن الطبيعي أن يزيد الخطأ في الاختبار الصعب المطبق على مجموعة من أصحاب القدرات المنخفضة، ويقل الخطأ إذا طبق على أصحاب القدرات المرتفعة.

رابعا :

تتطلب النظرية التقليدية نماذج متكافئة تماما للاختبار الواحد، وهذا مطلب صعب في التطبيق العملي مما يقلل من قيمة النتائج المبنية على نظرية القياس التقليدية. وقد استثارت تلك المشكلات العلماء المتخصصين في القياس النفسي والتربوي للبحث عن الدقة والموضوعية في القياس السلوكي حتى يقترب هذا القياس من القياس في العلوم الطبيعية، والتي تتميز بعدم تأثر نتائج القياس بالأداة المستخدمة طالما أنها أداة مناسبة لتقدير الظاهرة، كما يكون تدرج الأداة بوحدات قياس متساوية لا تعتمد ولا تتأثر بالعناصر التي تقدر عندها الظاهرة . ولذلك فان المقصود بالموضوعية عدم اعتماد درجة الفرد في الاختبار على عينة الأفراد، ففي الاختبارات النفسية المقننة تشتق الدرجات الخام من عينة التقنين ثم يوازن أداء الفرد الذي يطبق عليه الاختبار فيما بعد بمعايير مشتقة من هذه العينة، فإذا تغير العينة تفقد هذه المعايير دلالتها، أي يصبح الاختبار محكوماً بالعينة كما أن الفرد يحصل على الدرجة نفسها في أي اختبارين يقيسان نفس السمة،إلا أن درجة الفرد تختلف عادة باختلاف الاختبار الذي يطبق عليه، أي أن درجته تصبح محكومة بعينة المفردات التي يختبر بها

والقياس الموضوعي لا يعتمد في نظامه المرجعي على مقارنة نتائج أداء المتعلم بأداء الجماعة، أو بأداء الفرد ذاته في وقت لاحق أو في اختبار أو مقياس آخر، بل يعتمد على تقدير علاقة احتمالية بين الأداء الملاحظ للفرد في المقياس والسمات أو القدرات التي تكمن وراء هذا الأداء وتفسره

ميزات النظرية الكلاسيكية:

- صغر حجم العينة التي تحتاجها.
- تحليلاتها الرياضية بسيطة مقارنة مع النظرية الحديثة (IRT).
- حساب المؤشر (الإحصائي) للفقرة أو الاختبار يكون مباشر.
- التحليل لا يتطلب شروطا قاسية لمناسبة النموذج للبيانات

تحليل الفقرات حسب النظرية الكلاسيكية

- تطبيق وتحديد مؤشرات الفقرة وذلك بخطوات رياضية و حجم عينة ممثل للمجتمع.
- حذف الفقرات التي تعتمد على إحصائيات محكية وتحليل الفقرات المعيارية والذي يتضمن الصعوبة والتميز
- تعتمد مؤشرات الفقرة على خصائص العينة لذا يصبح من الضروري أن تكون العينة ممثلة

وحيث أن العينات غير المتجانسة تعطي نتائج عالية في تميز الفقرة وتحسب عن طريق معامل بوينت بايسيريال (P.Bis)أو معامل الارتباط بايسيريال(Bis) بينما ترتفع الصعوبة مع قدرة المجموعة العالية وينخفض مع متوسط المجموعة المنخفضة. والبحث عن الفقرات الضعيفة يتم من خلال دراسة إحصائيات الفقرة،والفقرة الضعيفة تحدد بمعامل الصعوبة من حيث قيمته العالية أو المنخفضة أو ارتباطا ضعيفا في العلامة الكلية على الفقرة، ومن المناسب الإشارة أن التحليل للفقرات في هذه النظرية مجتمعة لديها القدرة في تزويد مطور الاختبارات بمعلومات عن نوعية فقرات الاختبار بغض النظر عن النموذج المستخدم.



٢. اختيار الفقرات items selection

عند تطوير الاختبار فبالإضافة إلى الاهتمام بالصدق المتعلق بالمحتوى تختار الفقرات اعتمادا على خاصيتي صعوبة الفقرة وتمييزها والتوجه لاختيار الفقرة عالية التميز واختيار الفقرات بصعوبة معينة يحدده غرض الاختبار وتوزيع قدرات المجموعة التي سيقدم لها الاختبار فمثلا عند فحص الطلاب لمنحهم منحة دراسية يكون الاختبار صعبا على الجميع أما الاختبارات التحصيلية المعيارية المرجع تكون عالية التميز

النظرية الحديثة ونماذجها :. (Item Response Theory)

هي نظرية تدور حول الفقرة والأداء على الاختبار وكيفية ارتباط الأداء بالقدرة التي تقاس بالفقرات واستجابة الفقرة يمكن أن تكون متصلة أو أن متصلة أو مفرطة بمعنى أن علامات الفقرات يمكن أن تترتب ويمكن وجود قدرة واحدة أو أكثر حسب الأداء على الاختبار، حيث أن كثيرا من النماذج بنيت وطبقت على بيانات حقيقية بحيث تتضمن الخصائص التالية

- افتراض قدرة واحدة لتأكيد الأداء على الاختبار
- يمكن تطبيقها على بيانات مستقلة .
- الارتباط بين الأداء والقدرة يعبر عنه بمؤشر أو اثنين أو ثلاثة .

ومن المهم أن نعرف بان مظاهر الثبات في نظرية الاستجابة على الفقرة تعتبر مرتفعة وذلك لأنها تربط بين القدرة والإجابة على الفقرة،ولان إحصائيات الفقرة تقع على مقياس للقدرة وهذا ما لم يتوفر في النظرية التقليدية وقد جاء معنى الثبات من معرفة ما تقيسه الفقرة بطريقة افصل ومن معرفة العلاقة بين الأداء والقدرة. والذي يعرف على انه مجموع اقترانات معلومات الفقرات المكونة للاختبار والذي يمكننا من التنبؤ بعلامات المفحوصين عند قدرة محددة،فإذا تضمن الاختبار فقرات صعبة نسبيا فان المنحنى سوف يزاح إلى اليمين بمعنى أن العلامات سوف تصبح اقل ولذا فان هذا الاقتران يمكننا من توضيح كيفية اختلاف أداء المفحوصين من نفس القدرة وعلى اختبارين يقيسان نفس القدرة.إن هذا الاختلاف جزء من أخطاء

الدرجات السابقة .وفي الحقيقة فان هذا الاقتران يربط بين درجات القدرة في النظرية الحديثة والدرجة الحقيقية في النظرية التقليدية وذلك لان الدرجة المتوقعة على الاختبار عند تحديد مستوى القدرة وحسب التعريف هي درجة الطالب الحقيقية على فقرات الاختبار.

ومن الخصائص الأخرى للنظرية الحديثة، اقتران معلومات الفقرة Item Information Function والذي يوضح أكثر الفقرات مساهمة في قيمة القدرة حيث إن الفقرات ذات التمييز الأعلى عادة ما تكون أكثر مساهمة في قياس القدرة من الفقرات ذات التمييز المنخفض

ومن الخصائص الأخرى اقتران المعلومات للاختبار أيضا Test Information Function والذي يرمز له بالرمز (Φ) والذي يعتبر اقترانا معبرا عن مجموع اقترانات معلومات الفقرات التي يتكون منها الاختبار والذي من خلاله يمكننا معرفة وتحديد الخطأ المرتبط بقياس القدرة والذي يعني أن الخطأ يقل كلما زادت قيمة المعلومات التي نحصل عليها من الاختبار

تطبيقات نظرية السمات

إن نظرية استجابة الفقرة وفرت مجالا خصبا لتحسين تطبيقات قياس نفسية معروفة بالإضافة إلى مجالات تطبيق أخرى ومنها

- تقنين وتحليل الاختبارات : أصبح من الممكن مقارنة قيم القدرات الفردية عند التقنين بدلا من الاعتماد على التوزيع التكراري للعينة التي قنن عليها الاختبار اي أنه أصبح من الممكن إعداد بنوك من الأسئلة التي حددت مؤشراتها والاستفادة منها في تقنين الاختبارات.
- معايرة الاختبارات Equating أي المساواة بينها حيث يتم في كثير من الأحيان مقارنة شخصين أو أكثر بناء على درجاتهم الخام في نماذج مختلفة من اختبار معين، وهذا فيه شيء من الظلم والإجحاف، ولهذا فإن النظرية الحديثة توفر أسلوبا أفضل للمقارنة فهي لا توازن الدرجات الخام وإنما يكون الهدف

هو معرفة موقع القيمة الرقمية لقدرة الفرد على مقياس القدرة

- بناء الاختبارات. ويمكن تحديد نوع الاختبار الذي نريده (اختبار لاختيار أصحاب الكفايات العالية لممارسة مهمة خطيرة كالطيران مثلا) عند مستوى القدرة التي تم تحديدها سلفا
- الاختبارات التكيفية أو المفصلة . وهي اختبارات فردية يعطى فيها الفرد الأسئلة المناسبة لقدراته
- الكشف عن التحيز في الاختبار (test bias) فقد وفرت نظرية استجابة الفقرة نموذجا أفضل للكشف عن تحيز الاختبار

تحليل الفقرات حسب النظرية الحديثة

يتم تحليل فقرات الاختبار حسب النظرية الحديثة عادة وفقا للإجراءات التالية.

- تحديد معالم العينة الثابتة باستخدام طرق رياضية معقدة وعينة كبيرة الحجم.
- استخدام النموذج المطابق لاكتشاف الفقرات التي لا تتلاءم مع النموذج .

ويقصد بثبات العينة (sample invariance) أن مطوري الاختبارات ليسو بحاجة إلى عينة للمجتمع ولكنهم بحاجة إلى عينة غير متجانسة لحساب القيمة،وعلى أية حال فان تطوير الاختبار المستخدم يواجهه مشكلات كثيرة مثل كبر حجم العينة لحساب معالم الفقرات والبحث عن عقدان طفيف يختلف عن الطريقة في النظرية التقليدية، حيث أن تقييم الفقرات ما يتم بمدى ملائمتها للنموذج باستخدام اختبارات إحصائية أو تحليل (residuals) أي البواقي ومن المهم أن نشير إلى أن النموذج الملائم للبيانات هو أساسي لنجاح عملية التحليل فالفقرات الضعيفة يمكن توضيحها عن طريق التميز قيمة قليلة موجبة أو سالبة وعن طريق الصعوبة (ليست صعبة كثيرا وليست سهلة في نفس الوقت)

اختيار الفقرات :

يعتمد اختيار الفقرات حسب النظرية الكلاسيكية على طول الاختبار وفي النهاية يعتمد على المعلومات المساهمة في المعلومات الكلية التي يزودها الاختبار اقتران

معلومات الفقرة الذي يعطي مطور الاختبار نسبة مساهمة الفقرة في اقتران معلومات الاختبار وذلك باستقلالية عن الفقرات الأخرى في الاختبار وقد قدم(lord 77 birnbaum68) معلومات عن استخدام اقتران المعلومات في بناء الاختبار من خلال الإجراءات التالية .

١ ـ وصف شكل اقتران المعلومات المرغوبة بالنسبة إلى أي مدى للقدرات المرغوبة وهو ما يدعى target information fn

٢ ـ اختيار الفقرات حسب اقتران معلومات الفقرة الذي يملا المساحة تحت target information

٣ ـ بعد إضافة أو زيادة الفقرات للاختبار حسب اقتران معلومات الاختبار وفقرات الاختبار

٤ ـ الاستمرار في اختيار المعلومات إلى أن يقترب اقتران المعلومات الخاص بالاختبار من اقتران target information

فمثلا افترض أن أحد المطورين أراد أن يبني اختبار الخطوة الأولى يجب تعين اقتران المعلومات،ويجب إعادة حساب اقتران معلومات الفقرة بعد إضافة كل فقرة لتحديد المعلومات التي تحتاجها لإكمال الاختبار وبهذه الطريقة يكون لدينا فكرة جيدة عن قدرة المفحوصين و اختبار الفقرات

وخلاصة التميز بين النظريتين سنعرض للمقارنة التي أجراها راندال هامبلتون وجونز راسيل(Randall,Hambelton & j,Russel(1993 بين النظرية التقليدية ونظرية السمات الكامنة على النحو المبين في الجدول رقم (٢٠) التالي ٠

جدول (٢٠) مقارنة بين النظرية التقليدية ونظرية السمات الكامنة في القياس

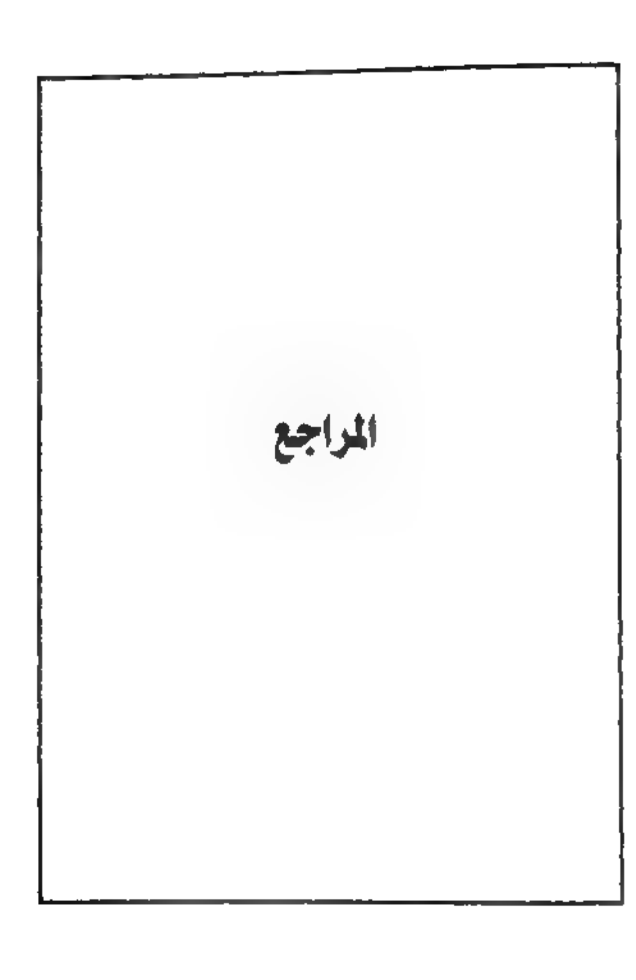
المعيار	النظرية التقليدية CTT	النظرية الحديثة IRT
النموذج	خطي	غير خطي
المستوى	اختبار	فقرات
الافتراضات	الافتراضات سهلة التطابق مع بيانات الاختبار	قوية صعبة المقابلة مع بيانات الاختبار
العلاقة بين القدرة - الفقرة	غير محددة	اقتران خصائص الفقرة
القدرة	علامة الاختبار و العلامة الحقيقية تحسب وتسجل على مقياس علامات الاختبار	علامات القدرة تسجل على مقياس من-∞ إلى∞
تباين الفقرات	إحصائيات الفرد غير مستقلة عن العينة	معالم الفقرات والأفراد مستقلة عن العينة إن كان النموذج ملائماً لبيانات الاختبار
إحصائية الفقرة	الصعوبة(p) والتميز (r)	A,b,c(الصعوبة التميز التخمين واقتران المعلومات)
حجم العينة	٢ ـــ ٥ بشكل عام	أكثر من ٥ بشكل عام
الدرجات	الدرجة الملاحظة = الدرجة الحقيقية + الخطأ القياس يعتمد على العينة وجود أخطاء كثيرة (الخطأ الثابت، خطأ القياس ، خطأ الصدفة أو العشوائية)	Rasch الدرجة = اللوجيت +/- المتبقي ، حيث أن اللوجيت = القدرة - صعوبة الفقرة أو IRT سيجما +/- الخطأ ، حيث إن سيجما تقدير القدرة على أساس الصعوبة ، والتمييز ، والتخمين القياس متحرر من العينة والفقرات
أخطاء القياس	خطأ القياس = الفرق بين الدرجة الملاحظة والدرجة الحقيقية تفسير الدرجة : X+/-(SEM)	خطأ القياس = الفروق بين الاستجابة الملاحظة والمتنبأ بها البواقي نموذج راش اللوجيت +/- (المتبقي) أوIRT · سيجما +/- (الخطأ)

الدقة في القياس	أن متوسط الدرجات التي تعود إلى الخطأ العشوائي =صفر معامل الارتباط بين الدرجات الحقيقية ودرجات الأخطاء العشوائية = صفر معامل الارتباط بين أي مجموعتين من درجات الأخطاء العشوائية = صفر	أحادية القياس الاستقلال المركزي للإستجابات تحرر القياس من الاختبار تحرر القياس من العينة التحرر من السرعة
	اقتران مميزة للفقرة ولا تختلف باختلاف الفقرات ويمثل بالمنحنى الاعتدالي التراكمي كمتغير عشوائي	اقتران معلومات خاص بكل فقرة من الفقرات ويمثل بشكل المنحنى اللوغاريتمي التراكمي

وبهذا يتبين نقاط الالتقاء وهي كثيرة بين النظريتين التقليدية والحديثة،وكذلك نقاط الاختلاف، هذا من شانه أن يبين حركة التطور التي حدثت في حقل القياس،من حيث المفاهيم والاعتبارات وبالتالي الافتراضات وتبعا لذلك الفوائد والتسهيلات التي طرأت على القياس النفسي وتحديدا في مما انعكس بالإيجاب على القياس النفسي وخاصة حقل التعليم والتعلم وفي النهاية ارجوا ا ن أكون قد قدمت مادة مفيدة للقارئين والمهتمين بالقياس النفسي والتربوي من خلال هذا الكتاب

تم بحمد الله





المراجع

المراجع العربية:

١ أبو جلالة، صبحي (١٩٩٩) اتجاهات معاصرة في التقويم التربوي وبناء الاختبارات وبنوك الأسئلة، مكتبة الفلاح للنشر الكويت.

٢ فؤاد أبو حطب و آمال صادق (١٩٩٦) "مناهج البحث وطرق التحليل الإحصائي في العلوم النفسية والتربوية والاجتماعية" ط ٢، القاهرة، الانجلو المصرية

٣ أبو علام، رجاء. (٢٠٠١) النظريات الحديثة في القياس والتقويم وتطوير نظم الاختبارات، المركز القومي للاختبارات والتقويم التربوي، بحوث المؤتمر الأول. الاختبارات والتقويم التربوي- رؤية مستقبلية.القاهرة

٤ أحلام، الشربيني ورضا حجازي. (٢٠٠٢) تقويم تحصيل تلاميذ مرحلة التعليم الأساسي في العلوم باستخدام نظرية السمات الكامنة بحوث المؤتمر العلمي السادس للتربية العلمية، المجلد الثاني، ٦١١-٦٤٨

٥ احمد الطيب (١٩٩٩) "التقويم والقياس النفسي والتربوي" الإسكندرية، المكتب الجامعي الحديث .

٦. أحمد عودة (١٩٩٢) مدى التوافق بين نموذج راش والمؤشرات التقليدية في اختيار فقرات مقياس اتجاه سباعي التدريج، مجلة كلية التربية بجامعة الإمارات، العدد (٨) يونيو، ١٥٣-١٧٩ .

٧ الدسوقي، عصام (١٩٩٨) "مدى فاعلية نموذج الحروف في تحديد المستوى لاختبار محكي المرجع"، مجلة كلية التربية جامعة المنصورة، العدد ٣٦، يناير، ص ص ٤٣ - ٧٣ .

٨. الدمنهوري، ناجي محمد قاسم (٢٠٠٠): فعالية استخدام كل من إستراتيجية التعلم التعاوني والتنافسي في التحصيل الدراسي والاتجاه نحو دراسة مادة

الرياضيات لدى طلاب المدرسة الثانوية الصناعية "مجلة كلية الآداب، جامعة المنوفية، العدد ٤١، ص ص ١ — ٤٥

٩. إسماعيل الوليلي (٢٠٠١). دراسة سيكومترية مقارنة لبعض نماذج الاستجابة للمفردة في انتقاء مفردات الاختبارات مرجعية المحك رسالة دكتوراه، جامعة الأزهر، كلية التربية

١٠. السيد أبو هاشم (٢٠٠٤) الدليل الإحصائي في تحليل البيانات باستخدام SPSS الرياض، مكتبة الرشد .

١١. إلياس، فوزي.(١٩٩٣) مشروع بنوك المفردات الاختبارية: مسار تربوي جديد في سلطنة عمان، رسالة التربية، عمان.

١٢. سعد عبد الرحمن (١٩٩٨). القياس النفسي (النظرية والتطبيق) ط٣، القاهرة، دار الفكر العربي .

١٣. الشايب، عبد الحافظ (2007). تقدير ثبات علامات عينة من المواد في جامعة آل البيت، مجلة جامعة دمشق – المجلد - 23 العدد الثاني .

١٤. شحتة عبد المولى (١٩٩٩) تقويم بناء الاختبارات المرجعة إلى المحك / المعيار في ضوء نظرية الاستجابة للمفردة والنظرية التقليدية رسالة دكتوراه، جامعة عين شمس، كلية التربية.

١٥. شريفين، نضال، (٢٠٠٣)، مدى تحقق معايير الفاعلية في معادلة اختبارين احدهما ثنائي التدريج والآخر متعدد التدريج وفق نماذج النظرية الكلاسيكية والنظرية الحديثة في القياس، رسالة دكتوراه غير منشورة، جامعة عمان العربية للدراسات العليا، عمان، الاردن.

١٦. شكري، سوميه (٢٠٠٦): فاعلية بعض طرق تقدير درجة الفصل في التنبؤ بالتحصيل اللاحق في الهندسة للمرحلة الإعدادية "رسالة ماجستير غير منشورة، كلية التربية، جامعة المنيا .



١٧. صلاح، مراد وأمين سليمان (٢٠٠٢)، الاختبارات والمقاييس في العلوم النفسية والتربوية – خطوات إعدادها وخصائصها . القاهرة، دار الكتاب الحديث

١٨. طومان، منار احمد (٢٠٠٣) "طرق حساب معامل ثبات الاختبار المرجع إلى المحك، دراسة إحصائية مقارنة"، رسالة ماجستير غير منشورة، معهد الدراسات والبحوث، جامعة القاهرة .

١٩. الطر يري، عبد الرحمن (١٩٩٦) الخصائص السيكومترية لاختبار الذكاء الإعدادي باستخدام نموذج راش مجلة دراسات نفسية، العدد (٤) أكتوبر، ٤٥٧-٤٧٣ .

٢٠. الطر يري، عبد الرحمن (١٩٩٧) القياس النفسي والتربوي ،نظريته، أسسه، تطبيقاته . الرياض، مكتبة الرشد.

٢١. عبابنه، عماد (٢٠٠٤). أثر حجم العينة وطريقة انتقائها وحده الفقرات وطريقة انتقائها على دقة تقدير معالم الفقرة والقدرة لاختبار قدرة عقلية باستخدام نظرية الاستجابة للفقرة رسالة دكتوراه، جامعة.

٢٢. عبد المسيح، عماد يوسف (١٩٨٢) "دراسة لبناء مقياس مرجعي الميز ن في مادة الطبيعة بالصف الأول الثانوي"، رسالة ماجستير غير منشورة، كلية التربية جامعة المنيا .

٢٣. عبد المسيح، عماد يوسف (١٩٩١)، استخدام نموذج راش اللوغاريتمي أحادي البارامتر في تحليل مفردات الاختبارات المعرفية مرجعية المعيار ثنائية القطب (دراسة تجريبية) . جامعة المنيا، كلية التربية، مجلة البحث في التربية وعلم النفس، العدد(٤) ابريل، ٤٤٣-٤٧٥

٢٤. علام، صلاح الدين محمود (١٩٨٥). "استخدام النموذج ذي الحدين في تقدير درجة القطع لاختبار محكي المرجع، دراسة إحصائية وتجريبية"، المجلة العربية للعلوم الإنسانية، تصدر عن جامعة الكويت، العدد ١٩، المجلد الخامس، ص ص ٢٦ – ٤٣

٢٥. علام، صلاح الدين (١٩٨٥)، تحليل بيانات الاختبارات العقلية باستخدام نموذج راش اللوغاريتمي الاحتمالي (دراسة تجريبية) . جامعة الكويت، المجلة العربية للعلوم الإنسانية، العدد(١٧)، ١٠٠-١٢٤

٢٦. علام، صلاح الدين محمود (١٩٨٦). "تطورات معاصرة في القياس النفسي والتربوي" إدارة التأليف والترجمة والنشر بجامعة الكويت، الكويت

٢٧. علام صلاح الدين محمود (١٩٨٧)، دراسة موازنة ناقدة لنماذج السمات الكامنة، والنماذج الكلاسيكية في القياس النفسي والتربوي جامعة الكويت، المجلة العربية للعلوم الإنسانية، العدد(٢٧)، ١٨-٤٤

٢٨. علام، صلاح الدين محمود (١٩٩١) "دراسة مقارنة لبعض طرق تحديد مستويات الأداء في اختبار مرجعي المحك"، المجلة المصرية للدراسات النفسية، تصدرها الجمعية المصرية للدراسات النفسية، ع١، سبتمبر، ص ص ٧٧ – ٩٦ .

٢٩. علام، صلاح الدين محمود (١٩٩٥) " الاختبارات التشخيصية مرجعية المحك في المجالات التربوية والنفسية والتدريبية"، القاهرة، دار الفكر العربي.

٣٠. علام صلاح الدين محمود (٢٠٠٠) القياس والتقويم التربوي والنفسي – أساسياته وتطبيقاته وتوجهاته المعاصرة ، القاهرة، دار الفكر العربي .

٣١. علام، صلاح الدين محمود(٢٠٠٠) .تطورات معاصرة في نظرية القياس الحديثة، ، جامعة الكويت .

٣٢. علام، صلاح الدين محمود(٢٠٠١) الاختبارات التشخيصية مرجعية المحك في المجالات التربوية والنفسية والتدريبية، القاهرة، دار الفكر العربي (٤٧٩-٥٢٢).

٣٣. غنيم، محمد ووليد القفاص (٢٠٠٠). إدراك الطلاب للمناخ الأكاديمي وعلاقته بمداخلهم للدراسة المجلة المصرية للدراسات النفسية، العدد (٢٥) يناير، ١٢٠ – ١٦٢

٣٤. كاظم، أمينة (١٩٨٨) استخدام نموذج راش في بناء اختبار تحصيلي في علم النفس وتحقيق التفسير الموضوعي للنتائج، جامعة الكويت

٣٥ الشرقاوي، أنور وآخرون (١٩٩٦) اتجاهات معاصرة في القياس والتقويم النفسي والتربوي القاهرة، الانجلو المصرية

٣٦ منسي، محمود عبد الحليم (٢٠٠٣) التقويم التربوي الإسكندرية، دار المعرفة الجامعية.

المراجع الأجنبية :

1. Alastarr& Hutchinson(1987). Calibrating graded assessment: rach partial credit analysis of performance in writing.Languege Testing.,(4)
2. Allen, M J &,Yen,W M (1979) Introduction to Measurement Theory Californu Cole publishing Company
3. Birnbaum, A "Some latent trait models and their use in inferring an examinee s ability " Part 5 in F M. Lord and M.R. Novick Statistical Theories of Mental Test Scores. Reading, MA Addison-Wesley, 1968.
4. Mableton, R.K., and Swaminathan, H(1984).. Item Response Theory: Principles and Applications.
5. R. K. Hambleton, H. Swaminathan, and H J.(1991). Fundamentals of Item Response Theory, Rogers, Sage.
6. Hambleton,R.& Jones R.(1993),Comparison of Classical test theory and item response theory and their application to test development .educational Measurement. Issues and practice(38-47).
7. Hans J eysenck (1980) Intelligence,Education,and the Genetic Model, john Wiley & sons ltd.
8. Hingham, MA. Kluwer, (1983). Item Response Theory: Application to Psychological Measurement. Nijhoff, Hulin, C. L., Drasgow, F., and Parsons, C.K.
9. Homewood, IL. Dow-Jones, Irwin: Lawrence Erlbaum, F. M. Lord(1980). Applications of Item Response Theory to Practical Testing Problems.
10. Gerhard(1980) .Individualized Testing on the basis of the Dichotomous Rasch Model. Fischer and Peter Pendle

11. Gronalund, N& Linn,R.(1980).Measurement and Evaluation in Teaching. New York, Macmillan publishing Co,inc.
12. Lord, F.M. ,(1980) Applications of Item Response Theory to Practical Testing Problems. Hillsdale, NJ Erlbaum.
13. Marcell Dekker, F.B. Baker(1992). Item Response Theory: Parameter Estimation Techniques
14. Mislevy, R.J , and Bock, R.D.(1986). PC-BILOG 3: Item Analysis and Test Scoring with Binary Logistic Models. Mooresville, IN. Scientific Software, Inc.
15. van der Linden W. J. and R. K. Hambleton, (1997).. Handbook of Modern Item Response Theory.
16. Wright, B.D., and Mead, R.J (1976). BICAL. Calibrating Items with the Rasch Model Research Memorandum No. 23 Statistical Laboratory, Department of Education, University of Chicago.
17. Wright, B D., and Stone, M.A.(1979). Best Test Design. Chicago:
18. Wright, D & ,Stone ,M.(1979).Best Test Design A handbook for Rash Measurement Chicago ,MESA press.





دار جرير
للنشر والتوزيع
www.darjareer.com


دار جرير
للنشر والتوزيع
www.darjareer.com





دار جرير
للنشر والتوزيع
عمان - شارع الملك حسين - مقابل مجمع الفحيص
هاتف 0096264651650 - فاكس 0096264643105
ص.ب 367 عمان 11118 الأردن
E-mail:dar_jareer@hotmail.com